

# Evaluatie uitbreidingsvraag en indicator toekomstige arbeidsmarktsituatie (ITA) tot 2008

Citation for published version (APA):

Bertrand-Cloodt, D. A. M. (2010). *Evaluatie uitbreidingsvraag en indicator toekomstige arbeidsmarktsituatie (ITA) tot 2008*. ROA. ROA Technical Reports No. 6  
<https://doi.org/10.26481/umarot.2010006>

## Document status and date:

Published: 01/01/2010

## DOI:

[10.26481/umarot.2010006](https://doi.org/10.26481/umarot.2010006)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Evaluatie uitbreidingsvraag en indicator toekomstige arbeidsmarktsituatie (ITA) tot 2008

Daniëlle Bertrand-Cloodt

## ROA Technical Report

ROA-TR-2010/6

**Research Centre for Education and the Labour Market**  
Maastricht University  
P.O. Box 616, 6200 MD Maastricht, The Netherlands  
T +31 43 3883647 F +31 43 3884914

[secretary-roa-sbe@maastrichtuniversity.nl](mailto:secretary-roa-sbe@maastrichtuniversity.nl)  
[www.roa.nl](http://www.roa.nl)

# **Evaluatie uitbreidingsvraag en indicator toekomstige arbeidsmarktsituatie (ITA) tot 2008**

Daniëlle Bertrand-Cloodt

ROA-TR-2010/6  
november 2010

**Research Centre for Education and the Labour Market**  
Maastricht University  
P.O. Box 616, 6200 MD Maastricht, The Netherlands  
T +31 43 3883647 F +31 43 3884914

secretary-roa-sbe@maastrichtuniversity.nl  
[www.roa.nl](http://www.roa.nl)

## **Voorwoord**

Deze evaluatiestudie heeft plaatsgevonden in het kader van het Project Onderwijs-Arbeidsmarkt (POA). Dit project wordt gefinancierd door het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW), het Uitvoeringsinstituut Werknemersverzekeringen (UWV), het UWV WERKbedrijf, het Ministerie van Landbouw, Natuur en Voedselkwaliteit (LNV), de samenwerkende kenniscentra voor beroepsonderwijs en bedrijfsleven COLO, Randstad Nederland en de Raad voor Werk en Inkomen (RWI). De auteur bedankt de leden van de begeleidingscommissie van het Project Onderwijs-Arbeidsmarkt voor het commentaar op een eerdere versie.

# 1 Inleiding

## 1.1 Uitgangspunten

Het Researchcentrum voor Onderwijs en Arbeidsmarkt (ROA) maakt in het kader van het Arbeidsmarktinformatiesysteem (AIS) iedere twee jaar middellange-termijnprognoses voor de arbeidsmarktperspectieven van opleidingen en beroepen. De laatste prognoses zijn terug te vinden in *De arbeidsmarkt naar opleiding en beroep tot 2014* (ROA, 2009) en het bijbehorende Arbeidsmarktinformatiesysteem (AIS). De prognoses uit het verleden worden regelmatig geëvalueerd.

In dit evaluatierapport zullen specifieke elementen van het prognosemodel worden geëvalueerd die in 2003 zijn gebruikt voor het opstellen van de uitbreidingsvraag en de indicator toekomstige arbeidsmarktsituatie (ITA) voor de periode 2003-2008 (ROA, 2003). De methodiek van de uitbreidingsvraag en de ITA zijn beschreven in *Methodiek arbeidsmarktprognoses en -indicatoren 2003-2008* (Cörvers et al., 2004). Bij de evaluatie in dit rapport wordt specifiek ingegaan op welke factoren de afwijkingen tussen de voorspelde en de gerealiseerde uitbreidingsvraag kunnen verklaren. Tevens zal de kwaliteit van de prognoses van de uitbreidingsvraag en de ITA voor de periode 2003-2008 zoveel mogelijk worden vergeleken met de kwaliteit van de eerdere prognoses en ITA, welke terug te vinden zijn in Dupuy (2005) en Dupuy (2009).

In de arbeidsmarktprognoses die in 2003 zijn samengesteld is voor het eerst gebruik gemaakt van een nieuw model voor de uitbreidingsvraag naar beroep.<sup>1</sup> Het model biedt een betere micro-economische onderbouwing en beter gebruik van

---

1. Zie F. Cörvers, A. Dupuy, S. Dijkman, B. Golsteyn, M. Hensen (2004), *Methodiek arbeidsmarktprognoses en -indicatoren 2003-2008*, ROA-W-2004/2, Universiteit Maastricht; F. Cörvers en A. Dupuy, *Explaining the Occupational Structure of Dutch Sectors of Industry, 1988-2003, 2006/7E*, Universiteit Maastricht. Over het destijds ontwikkelde model is inmiddels een artikel gepubliceerd in een wetenschappelijk tijdschrift: F. Cörvers en A. Dupuy (2010), 'Estimating Employment Dynamics across Occupations and Sectors of Industries', *Journal of Macroeconomics*.

sectorgegevens (kapitaal, R&D e.d.) dan het oude model. Het blijft echter onzeker in hoeverre het nieuwe model ook goed voorspelt. Nu de vijfjaarstermijn voor de prognoses van 2003 is afgelopen kunnen ze worden geëvalueerd volgens de gebruikelijke methodiek. Dit geeft op de eerste plaats inzicht in de kwaliteit en de betrouwbaarheid van de door het ROA gemaakte prognoses. Daarnaast kan de evaluatiestudie nieuwe inzichten geven in de bruikbaarheid van de gehanteerde prognosemethodiek. Naast een evaluatie van de uitbreidingsvraag naar beroep zal tevens de Indicator Toekomstig Arbeidsmarktperspectief (ITA) geëvalueerd worden.

Het evalueren van prognoses is met name van belang met het oog op toekomstige prognoseactiviteiten. Dat geldt niet alleen voor de opstellers van de prognoses maar ook voor de gebruikers. Voor de *gebruikers van de prognoses* over de toekomstige arbeidsmarktontwikkelingen is het nuttig informatie te hebben over de mate van betrouwbaarheid van de prognoses. In Borghans (1993) wordt getoond dat publieke voorspellingen een positieve invloed op de studiekeuze van leerlingen en daarmee op de werking van de arbeidsmarkt hebben, mits leerlingen een redelijk beeld hebben van de bruikbaarheid van deze prognoses. Voor het inzicht in de bruikbaarheid van de prognoses zijn twee zaken van belang. Ten eerste dient het voor gebruikers duidelijk te zijn op grond van welke argumenten een prognose tot stand is gekomen. Het totaalbeeld dat een prognose schetst dient verbijzonderd te worden naar de componenten waaruit zij is opgebouwd, zodat duidelijk wordt op grond waarvan bepaalde ontwikkelingen verwacht worden. Dit maakt het mogelijk de prognoses te vergelijken met de eigen verwachtingen over de toekomstige arbeidsmarktontwikkelingen en/of diverse andere informatiebronnen. Ten tweede is het van belang dat gebruikers een beeld hebben van de gemiddelde trefzekerheid van de voorspellingen, omdat dit mede bepaalt in welke mate zij rekening zouden moeten houden met de prognoses van het informatiesysteem. Om aan deze eis te voldoen is het derhalve belangrijk na te gaan op welke punten de prognoses redelijk

trefzeker zijn, en op welke punten de onzekerheden liggen. Ook moet bekeken worden op welke manier de mate van onzekerheid tot uitdrukking komt in de wijze waarop de prognoses worden gepubliceerd.

De evaluatie van de prognoses geeft informatie die benut kan worden voor een verdere verbetering van de prognosemodellen. Om deze reden is een goede evaluatie van de prognoses ook voor de *opstellers van de prognoses* van groot belang. Bij het opstellen van prognoses wordt op basis van inzichten in het functioneren van de arbeidsmarkt een keuze gemaakt tussen de vele mogelijke manieren om de arbeidsmarkt te modelleren. Indien alleen de kwaliteit van de data de kwaliteit van de prognoses zou bepalen, zou de enige les die uit de evaluatie getrokken kan worden, een roep om meer of betere data zijn. Een evaluatie van de prognoses kan echter ook nieuwe inzichten geven over de bruikbaarheid van de gehanteerde methode. Bovendien kan worden nagegaan of de aanpassingen in de methodiek die op basis van eerdere evaluatiestudies hebben plaatsgevonden ook in de verwachte verbeteringen hebben geresulteerd.

Naast de opstellers en de gebruikers hebben vanzelfsprekend ook financiers belang bij een evaluatie van de voorspellingen. Aan de ene kant kan deze evaluatie informatie geven omtrent de prioriteiten die gelegd moeten worden bij de verdere ontwikkeling van het informatiesysteem. Aan de andere kant is het voor de opdrachtgevers belangrijk te weten in hoeverre de prognoses bruikbaar zijn voor de door hen beoogde gebruiksdoelen.

## 1.2 Doel en opzet prognoses

De prognosegegevens en de overige arbeidsmarktgegevens die in het Arbeidsmarktinformatiesysteem (AIS) worden opgenomen zijn bedoeld voor zowel studiekeizers als voor andere actoren op de arbeidsmarkt zoals beleidsmakers (de overheid, de arbeidsvoorzieningsorganisatie, de sociale partners en het onderwijsveld) en het bedrijfsleven. De uitbreidingsvraagprognoses en de ITA

vormen dus een deel van alle informatie die beschikbaar is voor deze actoren. De doelstelling om informatie te genereren die bruikbaar is voor de studie- en beroepskeuzevoorlichting betekent dat de prognoses betrekking moeten hebben op aspecten die van belang kunnen zijn voor de studie- of beroepskeuzebeslissingen. Dat wil zeggen dat de prognoses inzicht moeten geven in de situatie op de arbeidsmarkt die een leerling aan zal treffen vanaf het moment dat deze, na zijn/haar studie te hebben voltooid, tot de arbeidsmarkt toetreedt. Deze voorwaarde legt dus eisen op aan de periode waarop de prognoses betrekking hebben en aan de groep op de arbeidsmarkt waarvoor ze relevant zijn. De einddatum van de prognoseperiode dient dus te vallen in de tijd dat de schoolverlaters de arbeidsmarkt zullen betreden en de verwachte perspectieven dienen op de toekomstige arbeidsmarktsituatie van de nieuwkomers gebaseerd te zijn. Echter ook voor werkgevers en beleidsmakers is dergelijke informatie over de middellange termijn van belang bij bijvoorbeeld strategische beslissingen over het recruitersbeleid, het HRM-beleid, en het onderwijs-, arbeidsmarkt- en technologiebeleid.

De prognoseresultaten dienen zoveel mogelijk geformuleerd te worden in algemeen gehanteerde begrippen en zo min mogelijk statistisch of economisch jargon te bevatten. Met name met betrekking tot statistische uitspraken over de betrouwbaarheid van de prognoses is een vertaling van groot belang. Bepaalde meer algemene beschrijvingen van de arbeidsmarkt kunnen voor beleidsmakers zeer relevant zijn, maar zijn voor individuen pas bruikbaar na een vertaling naar het individueel niveau. Zo is de uitbreidingsvraag voor een bepaalde opleidingscategorie vanuit een beleidsoogpunt interessant, maar is het voor een leerling belangrijker om te weten wat hiervan de consequentie is voor zijn of haar individuele kans op een bepaald soort werk na afloop van de studie.



Om de kwantitatieve gegevens van de uitbreidingsvraagprognoses en de ITA (en ook de andere prognoses, actuele data en indicatoren) toegankelijker te maken voor een bredere groep van gebruikers worden deze getransformeerd naar kwalitatieve typeringen. Op basis van de waargenomen waarden van de variabelen is telkens een classificatie gemaakt op een vijf-punts-schaal met de volgende typeringen 'erg laag', 'laag', 'gemiddeld', 'hoog' of 'erg hoog' (voor de uitbreidingsvraag) en 'zeer goed', 'goed', 'redelijk', 'matig' of 'slecht' (voor de ITA). De bedoeling van deze kwalitatieve typeringen is om de kwantitatieve gegevens beter toegankelijk te maken voor mensen die niet gewend zijn met dergelijke cijfers te werken. Ten eerste hoeft men zich nu niet te verdiepen in de meeteenheid van de variabele. Ten tweede krijgt men meteen een relatieve typering, zodat men geen inzicht hoeft te hebben in de spreiding van de variabele. Ten derde zorgt de verdeling in intervallen voor een typering die minder exact overkomt dan de cijfers zelf. Hiermee wordt de suggestie van nauwkeurigheid tot achter de komma vermeden en krijgen met name de arbeidsmarktprognoses een zekere bandbreedte.

Bij de empirische evaluatie moet rekening worden gehouden met de gezichtspunten van de gebruikers van de arbeidsmarktinformatie. Dit betekent dat het gehanteerde evaluatiecriterium, waarmee de omvang van de voorspelfouten wordt vastgesteld, vooral moet laten zien welke consequenties deze voorspelfouten hebben voor de individuele studie- en beroepskeuze of voor beslissingen van werkgevers of beleidsmakers. Bij een beoordeling van de prognoses vanuit een andere doelstelling zullen dan mogelijk ook andere eisen gesteld moeten worden aan zowel de inhoud en vorm van de informatie als aan het gehanteerde evaluatiecriterium. In hoofdstuk 2 wordt beargumenteerd dat een evaluatiecriterium vanuit de studie- en beroepskeuzeoptiek gebaseerd dient te zijn op procentuele voorspelfouten, terwijl het vanuit de beleidsoptiek wellicht meer voor de hand ligt om de absolute aantallen in de voorspelfout te gebruiken. De keuze van het evaluatiecriterium wordt verder besproken in paragraaf 2.2.

### 1.3 Opzet van het rapport

Zoals reeds eerder gezegd, wordt in dit rapport specifiek gekeken naar de uitbreidingsvraagprognoses en de ITA. In hoofdstuk 2 wordt de evaluatiesystematiek besproken. In hoofdstuk 3 wordt vervolgens nader ingegaan op hoe de uitbreidingsvraag tot stand is gekomen en wordt deze empirisch geëvalueerd. De kwaliteit van de uitbreidingsvraagprognoses zal daarbij worden vergeleken met de kwaliteit van de eerder gemaakte prognoses. Hetzelfde wordt gedaan voor de ITA in hoofdstuk 4. In hoofdstuk 5 worden ten slotte de belangrijkste bevindingen kort samengevat en wordt aangegeven welke aandachtspunten er uit deze evaluatiestudie volgen voor de verdere ontwikkeling van het informatiesysteem.

## 2 Een methode voor de empirische evaluatie

### 2.1 Inleiding

De methode van evalueren komt vrijwel overeen met de methode die is gehanteerd in de vorige evaluatierapporten (zie bijvoorbeeld Smits en Diephuis, 2001; Dupuy, 2005; Dupuy, 2009). Het uitgangspunt voor deze methode vormen de vragen die volgens Granger en Newbold (1986) een objectieve evaluatie van de uitbreidingsvraagprognoses moet beantwoorden:

- (1) Zijn de voorspellingen beter dan de beschikbare alternatieven?
- (2) Hoe 'goed' zijn de voorspellingen?
- (3) Kan de methode waarmee de voorspellingen zijn opgesteld zodanig aangepast worden dat een verbeterde voorspelkwaliteit verwacht mag worden?

In dit hoofdstuk komen alle onderdelen van de evaluatie kort aan de orde. Allereerst wordt in paragraaf 2.2 het gehanteerde evaluatiecriterium besproken. Dit

criterium geeft met name een waardering van het 'verlies' dat ontstaat door de voorspelfouten. Om meer zicht te krijgen op de oorzaken van de voorspelfouten worden vervolgens in paragraaf 2.3 enkele toetsen besproken. Er wordt achtereenvolgens ingegaan op de concentratie van voorspelfouten, mogelijke verklaringen voor de voorspelfouten en de over- en onderschatting van veranderingen. In paragraaf 2.4 wordt ingegaan op de evaluatie van kwalitatieve typeringen.

## 2.2 Het evaluatiecriterium

Uitgangspunt bij de beoordeling van de prognoses van het Arbeidsmarkt-informatiesysteem is, zoals gezegd, de doelstelling dat de prognoses geschikt moeten zijn voor de studie- en beroepskeuze. De voorspellingen van de ontwikkelingen van de prognoses zijn geformuleerd in aantallen personen. Voor een individuele leerling is echter niet de totale uitbreidingsvraag, het totale aantal werkzame personen in een beroep of een opleiding, of de absolute voorspelfout interessant, maar gaat het vooral om de relatieve voorspelfout:  $(x_i - \hat{x}_i) / y_i$ , waarin  $x_i$  staat voor de realisatie van een voorspelde grootte voor beroep of opleiding  $i$ , en  $\hat{x}_i$  de prognose voor dezelfde grootte aanduidt.  $y_i$  geeft het aantal werkzame personen in de voorspelde categorie aan. Onder de veronderstelling dat deze relatieve maatstaf normaal verdeeld is, is het zinvol om het kwadraat hiervan als *verlies* van de afzonderlijke prognose te beschouwen (Granger en Newbold, 1986).

$$V_i(\hat{x}_i) = \left( \frac{x_i - \hat{x}_i}{y_i} \right)^2 \quad (2.1)$$

Het individuele verliescriterium  $V_i$  geeft dus bij benadering het verlies dat relevant is voor de keuze van een individuele leerling, die eventueel het desbetreffende beroep of de desbetreffende opleiding zal kiezen. In dit rapport zal voor de uitbreidingsvraagprognose een tabel gepresenteerd worden, waarin prognoses, realisatie, voorspelfout en het verlies per beroepsgroep wordt weergegeven. Hoewel een dergelijk overzicht een eerste indruk van de kwaliteit van de voorspellingen

geeft en daardoor mogelijke problemen bij de prognoses aan het licht kan brengen, moet bij de interpretatie van dit overzicht gewaakt worden voor ad hoc redeneringen.

Om een beeld te kunnen krijgen van de totale kwaliteit van de prognoses, is het niet zo zinvol de prognoses voor iedere beroepsgroep of opleidingstype afzonderlijk in beschouwing te nemen, omdat veel voorspelfouten op zich een incidenteel karakter hebben. Het is moeilijk hieruit lessen te trekken. Doordat in het Arbeidsmarktinformatiesysteem prognoses zijn gemaakt voor 127 beroepsgroepen en 104 opleidingstypen<sup>2</sup>, is het gezien de hoeveelheid informatie dan ook wenselijk het verlies over alle beroepsgroepen of opleidingstypen te aggregeren. Op deze wijze kan beter gekeken worden naar de systematiek in de prognosefouten. Deze aggregatie maakt het mogelijk uitspraken te doen over de kansverdeling van het verlies over alle beroepsgroepen of opleidingstypen in plaats van een afzonderlijke beschouwing van iedere voorspelling te maken. De afzonderlijke verliezen per

- 
2. De uitbreidingsvraagprognoses worden in deze evaluatiestudie alleen geëvalueerd voor de beroepsgroepen en niet voor de opleidingstypen. De reden hiervoor is dat het voor het prognosemodel voor de uitbreidingsvraag naar opleidingstype erg moeilijk is geschikte verklarende variabelen te vinden. Het is doorgaans niet zo dat er slechts één bepaalde opleiding geschikt is voor een bepaald beroep. De Grip en Heijke (1991) laten zien dat het belangrijk is te onderkennen dat er voor het uitoefenen van een bepaald beroep meerdere opleidingstypen in aanmerking komen. De vraag naar arbeid per beroepsgroep kan in het algemeen dan ook door mensen met verschillende opleidingsachtergronden worden opgevuld. Aan de andere kant is er ook vanuit verschillende beroepsgroepen vraag naar een bepaald opleidingstype. De schaarsteverhoudingen op de arbeidsmarkt bepalen ook in hoeverre arbeidskrachten met een bepaalde opleidingsachtergrond uitwijken naar banen die onder het niveau liggen van de gevolgde opleiding. Borghans, van Eijs en de Grip (1994) wijzen er op dat dit betekent dat de voorspelde uitbreidingsvraag daardoor niet gelijk hoeft te zijn aan de feitelijke toekomstige werkgelegenheid per opleidingstype. Borghans, de Grip en Hoevenberg (1994) trekken daaruit de conclusie dat het niet zinvol is deze uitbreidingsvraagprognoses empirisch te evalueren. Daarom is het vooralsnog niet mogelijk een beeld te geven voor welke opleidingstypen de uitbreidingsvraagprognoses een slechte voorspelkracht hebben. In tegenstelling tot de uitbreidingsvraagprognoses wordt de indicator toekomstige arbeidsmarktsituatie in deze evaluatiestudie wél geëvalueerd voor de opleidingstypen.

beroepsgroepen of opleidingstypen kunnen worden geaggregeerd door op basis van het individuele verlies,  $V_i$ , het gemiddelde verlies te berekenen.

Bij het aggregeren van dit criterium moet er rekening mee worden gehouden dat er veel meer nieuwkomers op de arbeidsmarkt in een groot beroep of een grote opleidingscategorie terecht komen dan in een kleine. Door te wegen naar de omvang van het beroep of de opleidingscategorie wordt het *gemiddeld verlies* bepaald. Dit is het evaluatiecriterium op geaggregeerd niveau:

$$GV(\hat{x}) = \sum_i \frac{y_i}{y^{tot}} V_i(\hat{x}_i) = \sum_i \frac{y_i}{y^{tot}} \left( \frac{x_i - \hat{x}_i}{y_i} \right)^2 = \sum_i \frac{y_i}{y^{tot}} \frac{(x_i - \hat{x}_i)^2}{y_i^2} = \sum_i \frac{1}{y^{tot}} \frac{(x_i - \hat{x}_i)^2}{y_i} \quad (2.2)$$

Hierin is  $y^{tot} = \sum_i y_i$  de totale gerealiseerde omvang van de werkzame personen, terwijl het gemiddeld verlies (GV) het verlies van een gemiddelde leerling aangeeft.

De verliesfunctie geeft een schatting van de spreiding van de voorspellingen rond de realisatie en daarmee een antwoord op de tweede eis die Granger en Newbold (1986) stellen aan een objectieve evaluatie. Op grond van dit cijfer is het echter niet mogelijk een duidelijke uitspraak te doen over de kwaliteit van de prognoses. Er is immers geen informatie beschikbaar over wat een redelijke waarde voor het gemiddeld verlies zou zijn. Om te voorzien in een dergelijke beoordelingsmaatstaf kan de destijds gehanteerde prognose vergeleken worden met de voorspelkwaliteit van andere beschikbare prognoses, conform de eerste eis uit de lijst van Granger en Newbold (1986). Door de verhouding te nemen tussen de voorspelkwaliteit van de prognose en de voorspelkwaliteit van een referentievoorspelling, ontstaat een score die kleiner is dan 1 als de prognose beter is dan de referentieprognose en groter dan 1 als de prognose slechter is:

$$S(\hat{x}, x^{ref}) = \frac{GV(\hat{x})}{GV(x^{ref})} \quad (2.3)$$

Om invulling te kunnen geven aan de score-definitie, moet een bepaalde prognose als referentieprognose worden bestempeld. Evenals bij de vorige evaluaties van de ROA-prognoses (bijv. Smits en Diephuis, 2001; Dupuy, 2005; Dupuy, 2009) worden de uitbreidingsvraagprognoses vergeleken met de beschikbare arbeidsmarktinformatie in de situatie waarin gebruikers (onder wie leerlingen) die voor een bepaalde (studiekeuze)beslissing staan zouden verkeren indien ze geen prognoses uit het informatiesysteem tot hun beschikking zouden hebben (in het basisjaar van de prognoses). Deze referentieprognose wordt de “same as before (SAB) prognose” genoemd en is gelijk aan de situatie in het basisjaar (2002).

Een laatste punt van afweging is de schaal waarop de prognoses geëvalueerd worden. Veel beleidsmakers lijken met name te hechten aan de evaluatie van de absolute prognoseaantallen. In het kader van de studie- en beroepskeuze-voorlichting gaat het er echter niet alleen om een goede inschatting te maken van de absolute positie van een beroep of opleiding, maar vooral ook om het in beeld brengen van de relatieve positie. Daarom wordt de evaluatie in dit rapport doorgaans naast de *absolute prognoses* ook uitgevoerd op een prognose die gecorrigeerd is voor het totale volume-effect. Dat wil zeggen, de prognoses zijn vermenigvuldigd met een factor zodanig dat over alle categorieën geaggregeerd de prognoses gelijk zijn aan de realisatie. Deze prognoses worden aangeduid als de *relatieve prognoses*.

### 2.3 De oorzaak van voorspelfouten

Het gemiddeld verlies en de score die in de vorige paragraaf zijn besproken, geven informatie over de kwaliteit van de gemaakte voorspellingen. Op zich geven zij echter niet aan waardoor een prognose goed of slecht is uitgevallen. Op deze wijze wordt derhalve niet duidelijk hoe de prognosemethodiek op grond van de evaluatieresultaten kan worden aangepast. Dit laatste is volgens Granger en Newbold (1986) het derde belangrijke punt bij het uitvoeren van een objectieve

evaluatie. Om op basis van deze evaluatie conclusies te kunnen trekken die inzicht geven in de aard van de problemen die bij de prognoses verwacht kunnen worden en om consequenties met betrekking tot de gehanteerde methodiek te kunnen trekken, worden in aanvulling op de in de vorige paragraaf beschreven evaluatiemaatstaf, enkele analyses uitgevoerd die inzicht verschaffen in de oorzaken van de gemaakte voorspelfouten.

### *Verklaring van voorspelfouten*

Een overzicht van separate voorspelfouten kan bruikbaar zijn voor een subjectieve evaluatie. Derhalve worden in dit rapport tabellen met voorspelfouten voor elke beroepsgroep getoond. Een nadeel van het analyseren van de voorspelkwaliteit van de afzonderlijke beroepsgroepen is dat het systematische karakter van voorspelfouten niet aan het licht kunnen komen. Iedere voorspelfout wordt in dat geval als een incident op zich beschouwd waardoor de structuur achter deze voorspelfouten verloren gaat.

Het verschil tussen de voorspelling en werkelijke uitkomst kan het best worden gezien als de uitkomst van een stochastische variabele. Op het moment dat de voorspeller de prognose opstelt, is het niet mogelijk om te voorspellen hoe groot de voorspelfouten zullen zijn. Voor de evaluatie is het echter interessant om een idee te krijgen van de distributie van deze ‘random’ variabele. In de empirische evaluatie wordt derhalve een schatting gemaakt van de standaardafwijking van de relatieve voorspelfouten, onder de veronderstelling dat de voorspelfouten normaal verdeeld zijn:

$$\varepsilon_i = x_i - \hat{x}_i \sim N(\mu_i, \sigma_i^2) \quad (2.4)$$

met:

$$\sigma_i = (y_i)^\alpha e^{Z_i\beta} \quad (2.5)$$

$$\mu_i = \theta(\hat{x}_i - x_i^{\text{ref}}) \quad (2.6)$$

Hierin is  $x_i$  wederom de te voorspellen grootheid (uitbreidingsvraag per beroepsgroep) in het prognosejaar,  $\hat{x}_i$  de prognose van deze grootheid,  $x_i^{\text{ref}}$  de referentieprognose (i.e. de ‘naïeve’ prognose) en  $Z_i$  zijn de andere verklarende variabelen voor de standaardafwijking van de prognoses. De variabele  $y_i$  duidt op de omvang van de werkgelegenheid voor de beroepsgroep. Bij prognoses van de werkgelegenheid geldt dus dat  $y_i = x_i$ . Voor grotere beroepsgroepen kan verwacht worden dat de voorspelfout eveneens groter is.  $\sigma_i$  is een parameter voor de standaardafwijking van de voorspelfout. Vergelijking (2.5) is zo gespecificeerd dat deze standaardafwijking altijd positief is. In vergelijking (2.6) is  $\mu_i$  de verwachte waarde van de te voorspellen grootheid als er wordt gecorrigeerd voor een systematische over- of onderschatting van veranderingen. Op beide aspecten wordt hierna verder ingegaan.

### *Verklaring van de standaardafwijking*

Als eerste wordt met vergelijking (2.4) - (2.6) onderzocht of er verklarende variabelen zijn voor de standaardafwijking van de voorspelfout. De eerste factor in vergelijking (2.5), de grootte van de beroepsgroep  $y_i$ , is opgenomen om het schaaleffect te kunnen vaststellen. Er kan verwacht worden dat grotere beroepsgroepen ook te maken hebben met grotere voorspelfouten, maar waarschijnlijk zal deze voorspelfout niet volledig proportioneel toenemen.  $\alpha$  zal dus waarschijnlijk kleiner zijn dan 1 en groter dan 0. De keuze van de overige verklarende variabelen voor de omvang van de standaardafwijking van de



voorspelfouten ( $Z_i$ ) hangt af van de grootte die in beschouwing wordt genomen. Vergelijkbaar met de in de vorige paragraaf gemaakte indeling kan hierbij zowel gedacht worden aan inhoudelijke variabelen als aan variabelen die te maken hebben met de gehanteerde methodiek.

Het Arbeidsmarktinformatiesysteem bevat ook zogenaamde risico-indicatoren, namelijk de uitwijkmogelijkheden naar andere bedrijfsklassen en opleidingstypen, en de conjunctuurgevoeligheid van de werkgelegenheid. De uitwijkmogelijkheden geven bijvoorbeeld aan in welke mate de werkgelegenheid in een beroepsgroep gespreid is over verschillende bedrijfssectoren. Als deze spreiding groot is, zal een onverwachte verandering in een bepaalde bedrijfstak relatief weinig invloed hebben op de ontwikkelingen in die beroepsgroep. Verder geeft deze spreiding ook aan dat, indien er een tegenvallende vraag in één van de bedrijfstakken optreedt, er waarschijnlijk wel mogelijkheden zijn om in andere bedrijfssectoren werk te vinden. De conjunctuurgevoeligheid geeft aan in welke mate de werkgelegenheid van een beroepsgroep of opleidingstype meefluctueert met de conjunctuurgolven. Een hoge conjunctuurgevoeligheid betekent een grotere kans dat in de latere loopbaan de arbeidsmarktperspectieven op een gegeven moment verslechteren. Het ligt voor de hand te veronderstellen dat deze risico-indicatoren indirect ook een indicatie vormen voor de voorspelkwaliteit van de betreffende categorie. Deze indicatoren zijn mede bedoeld om een beeld te geven van de gevoeligheid van bepaalde beroepsgroepen of opleidingstypen voor exogene invloeden. Er kan dus verwacht worden dat er een samenhang bestaat tussen deze risico-indicatoren en de omvang van de voorspelfout. De methode op basis van (2.4) en (2.5) kan daardoor ook gebruikt worden om deze risico-indicatoren indirect te evalueren.

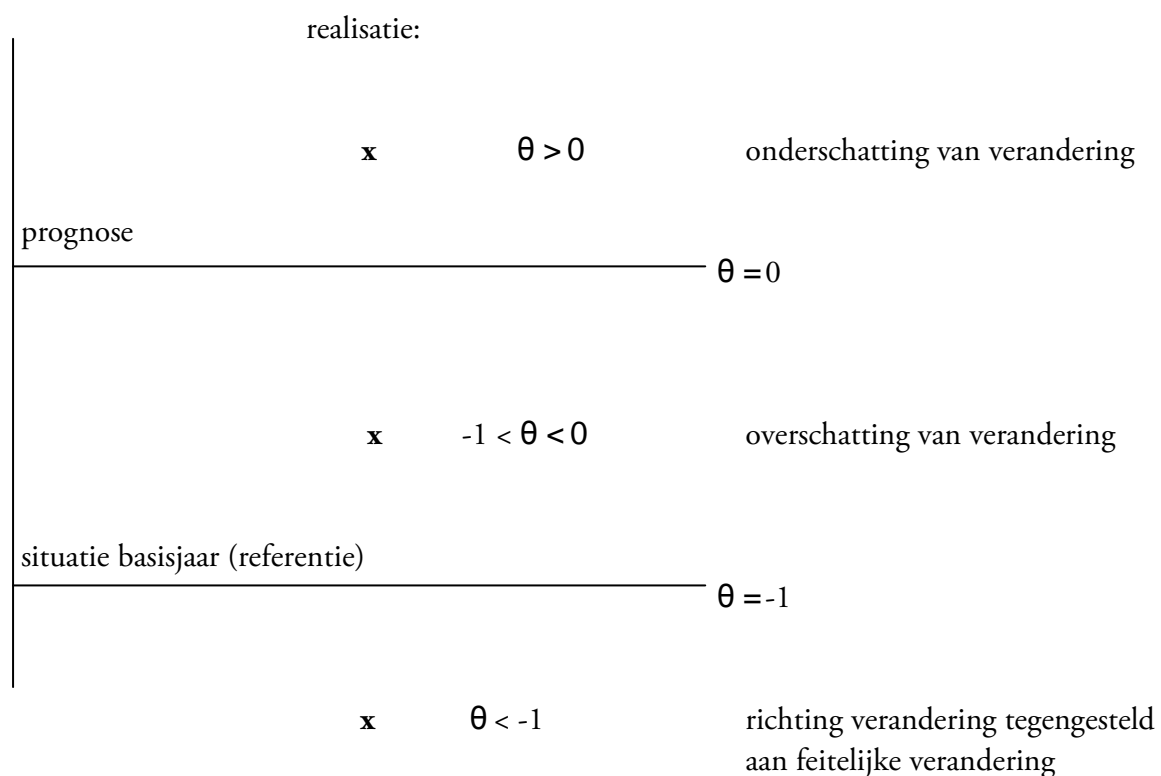
### *Over- of onderschatting van veranderingen*

Vergelijking (2.6) geeft een indicatie van de mate waarin er over- of onderschatting van de veranderingen heeft plaatsgevonden. Een belangrijke component in de

prognoses is veelal de trend in de te voorspellen grootheid. Een cruciaal punt bij het extrapoleren van bestaande trends is de wijze waarop deze extrapolatie plaats dient te vinden. Het zonder meer doortrekken van trends leidt vaak, zeker op de langere termijn, tot onwaarschijnlijke resultaten. In dat geval wordt soms de trend helemaal niet opgenomen in de prognose, of wordt er een aanpak gekozen waarbij de invloed van de trend afneemt in de loop van de tijd. Vergelijkbare problemen doen zich in principe voor bij iedere verklarende variabele. Als een schatting van een parameter toevallig hoog uitvalt, zal men, wanneer de prognose op deze hoge parameterwaarde wordt gebaseerd, de invloed van deze variabele overschatten. Derhalve wordt er vaak bij niet significante waarden van de parameterschatting gekozen voor het negeren van de parameter bij het opstellen van de prognose.

**Figuur 2.1**

De mogelijke waarden van de onderschattingscoëfficiënt  $\theta$  bij een over- of onderschatting van veranderingen



Een evaluatie is een geschikt middel om te bekijken in hoeverre de geschatte invloeden van de exogene variabelen in het verleden op een juiste manier zijn doorgetrokken naar de toekomst. Een te voorzichtige extrapolatie zal een onderschatting van verandering bewerkstelligen, terwijl een overdreven extrapolatie een overschatting van verandering is (Theil, 1958; Borghans, 1993).

Op basis van deze vergelijking is het mogelijk een verwachting van het verschil tussen de realisatie en de prognose te formuleren. De rechterkant van vergelijking (2.6) bevat alleen grootheden die bekend waren op het moment dat de prognose werd opgesteld. Daardoor zou het destijds in principe mogelijk zijn geweest een betere prognose op te stellen die deze over- of onderschatting van veranderingen niet bevatte. Figuur 2.1 geeft aan wat de betekenis is van bepaalde parameterwaarden van  $\theta$ . In de figuur is verondersteld dat de prognose hoger uitvalt dan de waarde van de grootheid die bekend was in het basisjaar (c.q. de referentieprognose). Het omgekeerde is echter ook mogelijk. Dan ontstaat een situatie die het spiegelbeeld is van de in figuur 2.1 geschetste situatie.

In de figuur zijn drie gebieden te onderscheiden met twee grenswaarden. Als  $\theta = 0$  dan is gemiddeld genomen de realisatie gelijk aan de prognose. De kwaliteit van de prognose kan dan weliswaar nog slecht zijn, maar er vindt geen systematische over- of onderschatting van de veranderingen plaats. Als  $\theta > 0$  dan ligt de realisatie van de voorspelde grootheid gemiddeld verder van de waarde in het basisjaar af dan de prognose. Er is dan dus een onderschatting van veranderingen. Omgekeerd betekent  $\theta < 0$  dat de realisatie gemiddeld dichterbij de waarde in het basisjaar ligt dan de prognose. Er is dan dus sprake van een overschatting van veranderingen. Hierbij wordt verondersteld dat  $\theta > -1$ . Als  $\theta = -1$  dan is de overschatting van de veranderingen totaal. Dat wil zeggen dat in dat geval iedere voorspelde verandering geen enkele informatieve waarde had en men net zo goed de referentieprognose als prognose had kunnen hanteren. Als  $\theta$  kleiner is dan  $-1$ , dan is er zelfs sprake van een omkeringseffect. Dit betekent dat daar waar dalingen werden voorspeld er

gemiddeld sprake was van een stijging, terwijl er bij voorspellingen van stijgingen gemiddeld een daling is opgetreden.

Een over- of onderschatting van veranderingen kan veroorzaakt worden door een te hoge of te lage gemiddelde prognose. Als relatieve prognoses worden geanalyseerd speelt dit echter geen rol meer. In dat geval betekent een overschatting van veranderingen dat de voorspelde groei gemiddeld lager uitviel dan verwacht, terwijl tegelijkertijd ook de voorspelde krimp minder negatief uitviel dan voorzien.

Een overschatting van veranderingen hoeft niet alleen veroorzaakt te zijn door het te ver doortrekken van trends, maar is vaak het gevolg van een grote storingscomponent in de prognoses. Als prognoses naast een informatief gedeelte ook ruis bevatten kan het zinvol zijn de voorspelde veranderingen gedeeltelijk te negeren. Dergelijke ruis wordt doorgaans veroorzaakt door meetfouten in de gebruikte data, bijvoorbeeld vanwege de steekproeffout. Des te groter deze storingscomponent des te voorzichtiger men zou moeten zijn met het hanteren van de prognoses. In vergelijking (2.6) wordt dit aangegeven door een waarde van  $\theta$  die dicht bij  $-1$  ligt.

De vergelijkingen (2.4) - (2.6) worden uiteindelijk door middel van een OLS-regressie geschat met daarin de logaritme van de relatieve voorspelfouten (voorspelfouten zijn zo gespecificeerd dat ze altijd positief zijn) als afhankelijke variabele en een indicator voor onderschatting, de risico-indicatoren (i.e. spreiding over sectoren en opleidingstypen, conjunctuurgevoeligheid van de werkgelegenheid), de logaritme van het aantal werkenden in een beroepsgroep en eventuele overige controlevariabelen als regressors.

## 2.4 De evaluatie van kwalitatieve typeringen

De geschetste evaluatiemethode had tot nog toe betrekking op de puntvoorspellingen die voortkomen uit de prognosemodellen van het informatiesysteem. De uiteindelijke presentatie van de prognoses van de uitbreidingsvraag en de ITA hebben echter plaatsgevonden door middel van een kwalitatieve typering van de prognoseresultaten. In deze evaluatiestudie worden ook deze kwalitatieve typeringen geëvalueerd. Hiertoe worden ontwikkelingen die feitelijk hebben plaatsgevonden volgens hetzelfde indelingsschema voorzien van een kwalitatieve typering, waarna kan worden nagegaan in hoeveel gevallen de oorspronkelijke typering overeenkomt met de realisatie (zie ook De Grip, Heijke en Berendsen, 1991). Een dergelijke evaluatie wordt gemaakt op basis van een matrix waarbij de oorspronkelijke typering en de realisatie tegen elkaar worden afgezet. In dit rapport wordt hierbij steeds uitgegaan van de relatieve prognoses. Opgemerkt dient echter te worden dat bij een dergelijke evaluatie gelijktijdig de voorspelkwaliteit van de prognoses als de wijze van typeren wordt geëvalueerd. Relatief slechte prognoses die getypeerd worden in kwalificaties die een breed interval vertegenwoordigen, kunnen toch heel bevredigend zijn geweest (Borghans, Van Eijs en De Grip, 1994). Omgekeerd zullen zeer nauwkeurige prognoses vaak een onjuiste typering krijgen als de kwalificaties gebaseerd zijn op extreem smalle intervallen.

## 3 De uitbreidingsvraag tot 2008

### 3.1 Inleiding

De eerste component van de vraagprognoses in het informatiesysteem onderwijs-arbeidsmarkt<sup>3</sup> betreft de uitbreidingsvraag. De uitbreidingsvraag varieert aanzienlijk

---

3. Voor meer informatie over de opzet van het prognosemodel van het informatiesysteem onderwijs-arbeidsmarkt, zie bijvoorbeeld Dupuy (2005).

tussen beroepsgroepen en opleidingstypen en is daarom moeilijker te voorspellen dan de andere componenten. Daarbij komt dat de tijdreeksen waarop de voorspellingen gebaseerd worden erg kort zijn. In dit hoofdstuk wordt allereerst in paragraaf 3.2 de gehanteerde methodiek besproken die ten grondslag lag aan de prognoses voor de periode 2003-2008. Vervolgens worden de prognoses in paragraaf 3.3 per sector en in paragraaf 3.4 per beroepsgroep empirisch geëvalueerd.

### 3.2 Prognosemethodiek

Figuur 3.1 geeft aan hoe de prognoses van de werkgelegenheidsontwikkeling voor sector, beroepen (en opleidingen) tot stand komen. Uitgangspunt voor de prognoses waren de werkgelegenheidsprognoses per bedrijfssector van het Centraal Planbureau (CPB). Vanuit de prognoses per bedrijfssector wordt een prognose opgesteld voor beroepssegmenten. Vervolgens vindt er vanuit de werkgelegenheidsprognoses per beroepssegment een verdere uitsplitsing plaats naar de beroepsgroepen. De beroepenindeling is gebaseerd op de Standaard Beroepen Classificatie uit 1992 (SBC 1992) van het CBS (1993). De beroepssegmenten zijn gebaseerd op de indeling op 2-digit-niveau; de beroepsgroepen op de indeling op 3-digit-niveau.<sup>4</sup> Er worden in totaal 43 beroepssegmenten en 127 beroepsgroepen onderscheiden.<sup>5</sup> De voorspelde uitbreidingsvraag per beroepssegment was tevens het uitgangspunt voor de prognoses van de uitbreidingsvraag per opleidingstype.

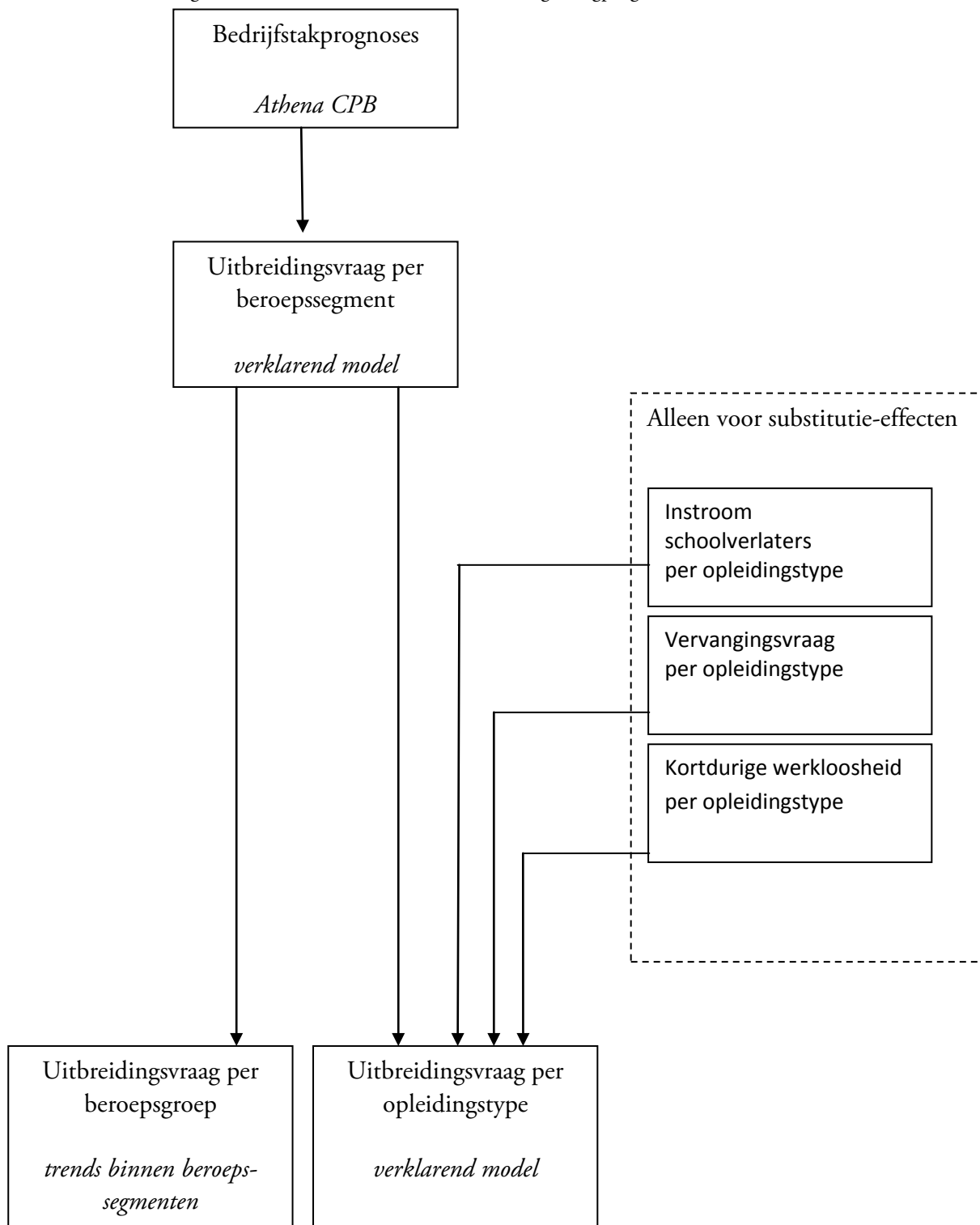
---

4. Bij de beroepsgroepindeling zijn de elementaire beroepen opgesplitst in zes beroepsgroepen, terwijl de middelbare procestechnische beroepen (code 471) verder zijn opgesplitst in de beroepsgroepen *procesoperators* en *bakkers en slaggers*.

5. Zie de 2-digitcodes en bijbehorende benamingen in de *ROA-classificatiegids* (ROA, 2002, tabel 6 en 7).

**Figuur 3.1**

Overzicht van de gehanteerde methodiek voor de uitbreidingsvraagprognoses



## *Bedrijfssectorprognoses*

Op basis van het Athena-model (CPB, 1993) maakt het CPB prognoses van de werkgelegenheidsontwikkeling per bedrijfssector. De sectorale werkgelegenheidsprognoses van het CPB hebben betrekking op het 'voorzichtige' scenario van de werkgelegenheidsramingen voor bedrijfssectoren in de publicatie *Economische Verkenning 2003 – 2006* van december 2001. Het CPB heeft tevens, in het verlengde van deze ramingen, een projectie gemaakt voor de verschillende bedrijfssectoren van 2006 tot 2010. In afwijking van deze middellangetermijnramingen voor de jaren tot 2008 is voor de bedrijfssectorprognoses van de uitbreidingsvraag voor 2003 en 2004 gebruik gemaakt van het *Centraal Economisch Plan 2003* (CEP) van het CPB. Door het gebruik van deze kortetermijnprognoses worden de gehanteerde middellangetermijnramingen gecorrigeerd met de meer recente inzichten in de groeiverwachtingen van het CEP.

De arbeidsvolume-prognoses van het CPB zijn met behulp van prognoses van de P/A-ratio's van het CPB omgerekend naar werkzame personen. De procentuele ontwikkelingen en de prognoses van het aantal werkenden in het informatiesysteem worden weergegeven ten opzichte van het aantal werkenden in het basisjaar 2002 volgens de Enquête Beroepsbevolking (EBB) van het CBS.

## *Het beroepenmodel*

In het beroepenmodel wordt de voorspelde werkgelegenheidsontwikkeling per bedrijfstak omgezet naar de werkgelegenheidsontwikkeling per beroepssegment en beroepsgroep. Hierbij wordt de veronderstelling gemaakt dat de werkgelegenheid per beroepsgroep volledig door de vraagzijde van de markt bepaald wordt. De ontwikkelingen in de vraag per beroepsgroep worden bepaald door de werkgelegenheidsverschuivingen tussen bedrijfssectoren en de veranderingen in de beroepenstructuur van de werkgelegenheid per bedrijfssector.



De uitbreidingsvraag naar beroep is geschat met EBB-data van 1987 tot en met 2002. Het gaat bij deze werkgelegenheidscijfers om werkzame personen die een werkverband van minstens 12 uren per week hebben. Het CBS is vanaf 1994 overgegaan op een nieuwe bedrijfssectorindeling. Om toch de EBB gegevens voor de periode 1987 tot en met 1993 te kunnen gebruiken is een schatting gemaakt van aantallen werkenden per bedrijfssector volgens de nieuwe indeling. Dit was mogelijk omdat in 1994 zowel de oude als de nieuwe classificatie gehanteerd zijn. De verhoudingen van de aantallen werkenden per beroepssegment per bedrijfssector volgens de oude en nieuwe indeling in dat jaar zijn dus bekend.

Verder zijn in de EBB-matrices beroepssegment x bedrijfssector de aantallen beneden de CBS-ondergrens gelijkgesteld aan 0. Hierdoor verdwijnen in iedere bedrijfssector een aantal kleinere beroepssegmenten. Beroepssegmenten die voor één of meerdere jaren niet voorkomen in de tijdreeks zijn weggelaten. In totaal zijn er 217 combinaties van beroepssegmenten en bedrijfssectoren waarvoor de werkgelegenheidsontwikkeling is geschat.

De beroepenstructuur binnen bedrijfssectoren is geschat aan de hand van een nieuw model van Cörvers en Dupuy (2003).<sup>6</sup> In dit model spelen, behalve de tijdtrend, vier factoren een rol bij de verklaring van de veranderingen in de vraag naar beroepen binnen bedrijfssectoren. De beroepenstructuur binnen bedrijfssectoren wordt bepaald door het productieniveau ('non-homothetic production function') per bedrijfssector, de kapitaalintensiteit per bedrijfssector ('capital-skill

---

6. Het oude model was gebaseerd op Borghans en Heijke (1994). Zij gebruikten een random-coëfficiënten-model met verklarende variabelen uit het Athena-model. Zij deden de aanbeveling om nader onderzoek te verrichten naar variabelen waarmee de werkgelegenheidsontwikkeling naar beroep beter verklaard en voorspeld kan worden. Een aangepaste versie van Cörvers en Dupuy (2003) is gepubliceerd in *Journal of Macroeconomics* (Cörvers en Dupuy, 2010). Dit betekent dat voor het schatten van de beroepssegmenten binnen bedrijfssectoren het nieuwe error-correctie-model van Cörvers en Dupuy (2003; 2010) gebruikt is, terwijl voor het schatten van de beroepsgroepen binnen segmenten het al eerder gebruikte random-coëfficiënten-model gebruikt is.

complementarity'), het gebruik van nieuwe meer kennisintensieve technologieën ('skill biased technological change') en de relatieve lonen op de verschillende arbeidsmarktsegmenten. Om het model te schatten zijn voor de 13 verschillende bedrijfssectoren data verzameld over de toegevoegde waarde, de investeringen in kapitaal, de loonsom en de investeringen in Research and Development (R&D) tussen 1987 en 2002. De benodigde data zijn ontleend aan de zogenaamde 'Lange Reeksen' van het CPB. Deze reeksen zijn gebaseerd op de Nationale Rekeningen van het CBS, en worden door het CPB gebruikt als data-input voor de schattingen van de vergelijkingen in het Athena-model (CPB, 1993). Alleen de R&D-gegevens zijn direct afkomstig van het CBS (via Statline), en zijn gebaseerd op de R&D- en innovatie-enquêtes van het CBS onder bedrijven, research-instellingen en universiteiten.

In de eerste stap van de schattingsprocedure is getoetst of de verschillende tijdreeksen al dan niet stationair zijn, dat wil zeggen er is gekeken naar de significantie van een stochastische of deterministische trend in de werkgelegenheidsontwikkeling van de beroepssegmenten per bedrijfssector en de bovengenoemde vier verklarende variabelen. De 'unit root'-testen wijzen uit dat er bijna altijd sprake is van een stochastische trend. Tevens is er in meer dan driekwart van de tijdreeksen een significante deterministische trend. Uit de resultaten van de empirische toetsing kan geconcludeerd worden dat alle tijdreeksen non-stationair en geïntegreerd van de eerste orde zijn. Derhalve wordt er getoetst of er sprake is van coïntegratie tussen de werkgelegenheidsaandelen van beroepssegmenten per bedrijfssector en de verklarende sectorvariabelen.

In de test op coïntegratie wordt gekeken naar de afwijking van het aandeel van de werkgelegenheid van een beroepssegment in een bedrijfssector van het geschatte langetermijnverband (zie de 'error correctie' term in vgl. 3.1). Het blijkt dat de residuen stationair zijn voor meer dan driekwart van de combinaties van

beroepssegment en bedrijfssector. Voor de werkgelegenheidsaandelen van beroepssegment  $j$  in bedrijfssector  $i$ , weergegeven door  $l_{ij,t}$ , worden daarom aan de hand van de betreffende tijdreeksen van 1987-2002 ( $t$ ) de volgende vergelijkingen geschat volgens de ‘error correctie’ specificatie zoals voorgesteld door Engle en Granger (1987):

$$\Delta l_{ij,t} = \alpha_{ij} + \sum_k \lambda_{kij} \Delta x_{ki,t} + \gamma_{ij} \sum_k (l_{ij,t-1} - \beta_{kij} x_{ki,t-1}) + \varepsilon_{ij,t}^* \quad (3.1)$$

waarbij  $x_{ki,t}$  de verklarende variabelen  $k$  per bedrijfssector  $i$  representeren, en  $\varepsilon_{ij,t}^*$  is een storingsterm.  $\alpha_{ij}$  en  $\lambda_{kij}$  zijn de parameters voor respectievelijk de constante en de  $k$  verklarende variabelen ( $k=4$ ). Bovendien geven de parameters  $\gamma_{ij}$  en  $\beta_{kij}$  de effecten weer van respectievelijk de afwijkingen van het langetermijnevenwicht (de ‘error correctie’) en de verklarende variabelen  $k$  op de werkgelegenheidsaandelen van beroepssegment  $j$  in bedrijfssector  $i$ . Het ‘error correctie’ mechanisme is in ongeveer 90% van de combinaties van bedrijfssectoren en beroepssegmenten statistisch significant.

Voor de prognose van de werkgelegenheidsaandelen van beroepssegmenten binnen bedrijfssectoren is gebruik gemaakt van de eerder genoemde Athena-prognoses van het CPB. Het betreft hier niet alleen de prognoses van de werkgelegenheid naar bedrijfstak, maar ook de bedrijfstakprognoses voor de toegevoegde waarde, de investeringen in kapitaal en de loonsom. De prognoses van de investeringen in R&D zijn gebaseerd op een extrapolatie van de tijdreeksen van het CBS. De bedrijfstakprognoses van de werkgelegenheid van het CPB (zie hierboven) en de prognoses van de veranderingen in de werkgelegenheidsaandelen van beroepssegmenten binnen bedrijfssectoren zijn toegepast op het aantal werkenden in 2002 volgens de EBB. De aggregatie over dezelfde beroepssegmenten binnen de verschillende bedrijfssectoren resulteert in de prognoses van de uitbreidingsvraag per beroepssegment.

Vervolgens zijn de prognoses per beroepssegment verbijzonderd naar beroepsgroepen. Hierbij is gebruik gemaakt van het random-coëfficiënten-model. De werkgelegenheidsgroei per beroepsgroep wordt daarbij geschat als afwijking van de totale groei van het beroepssegment waar de betreffende beroepsgroep onder valt. Om stabiele parameterschattingen te krijgen zonder daarbij de specificiteit van de afzonderlijke beroepen aan te tasten is gebruik gemaakt van een random-coëfficiënten-model.<sup>7</sup> In dit random-coëfficiënten-model worden de parameterwaarden bepaald als een gewogen gemiddelde van aan de ene kant een gemiddelde parameterwaarde over de beroepsgroepen waarover gepoold wordt, en aan de andere kant schattingen voor de afzonderlijke beroepsgroepen. Het gewicht wordt bepaald door de nauwkeurigheid van beide onderdelen. Als er veel variatie is tussen de verschillende beroepsgroepen binnen een beroepssegment is een gepoolde schatting relatief onnauwkeurig en komt er meer gewicht op de afzonderlijke schattingen te liggen. Als deze afzonderlijke schattingen echter een hoge standaardfout hebben wordt hun gewicht verlaagd. Op deze wijze ontstaat er een optimale combinatie van de informatie van de gepoolde gegevens en de afzonderlijke schattingen.

$$\Delta l_t^g = \Delta l_t^j + \beta_0^g + \beta_1^g (\Delta l_{t-1}^g - \Delta l_{t-1}^j) + \varepsilon_t \quad (3.2)$$

waarbij:

$l_t^g$  = logaritme van het aantal werkzame personen in beroepsgroep  $g$  op tijdstip  $t$ ;

$l_t^j$  = logaritme van het totaal aantal werkzame personen in beroepssegment  $j$  waartoe beroepsgroep  $g$  wordt gerekend op tijdstip  $t$ .

Overigens is bij de schattingen van dit model een dummy-variabele voor 1992 en 1996 opgenomen. In de periode 1992-1995 heeft het CBS niet het beroep gecodeerd van mensen die een aanstelling hadden van minder dan een half jaar

---

7. Borghans en Heijke (1994) geven een uitvoerige beschrijving van dit model.

zonder vooruitzicht op verlenging. Omdat dit tot 1992 wel gebeurde en vanaf 1996 ook weer neemt het aantal mensen van wie het beroep niet gecodeerd is van 1991 op 1992 sterk toe en van 1995 op 1996 weer sterk af. Deze breuk in de data wordt opgevangen door een dummyvariabele voor 1992 en 1996 aan vergelijking (3.2) toe te voegen. Daarbij is verondersteld dat voor elke beroepsgroep in een bedrijfssector de proportionele afname van 1991 op 1992 als gevolg van niet coderen gelijk is aan de proportionele toename in 1996.

### *Nederland in 2003*

De middellangetermijnprognoses van de uitbreidingsvraag die in dit rapport worden geëvalueerd zijn gemaakt in de zomer van 2003, en nemen als uitgangspunt de macro-economische verwachtingen tot 2008. In deze paragraaf wordt in het kort geschetst wat in 2003 de verwachte ontwikkeling op de arbeidsmarkt was voor de prognoseperiode, van 2003 t/m 2007 (zie ROA, 2003).

Evenals aan het begin van de jaren '80 en '90 gaat het in 2003 slecht met de Nederlandse economie. Er is nog nauwelijks economische groei en de werkgelegenheid krimpt. Wat betreft de arbeidsmarktontwikkeling springt de in hoog tempo oplopende werkloosheid het meest in het oog. Dit is een grote verandering ten opzichte van de grote krapte die er tot 2002 op de arbeidsmarkt heerste. Er wordt echter verwacht dat deze omslag niet op alle arbeidsmarktsegmenten in dezelfde mate merkbaar zal zijn. Zo verwacht het Centraal Planbureau (CPB) voor de overheidssector, waaronder ook de zorgsector, nog steeds een groeiende werkgelegenheid. Bovendien wordt verwacht dat de conjunctuur zich op de middellange termijn weer zal herstellen.

De verwachte daling in de werkgelegenheid tot en met 2004 wordt veroorzaakt door de neergaande conjunctuur. Er wordt een beperkte groei van het Bruto Binnenlands Product (BBP) verwacht van 1,0% in 2004. Na 2004 zal de

werkgelegenheidsgroei zich naar verwachting herstellen, waardoor per saldo een geringe gemiddelde jaarlijkse werkgelegenheidsgroei voor de periode tot 2008 resulteert.<sup>8</sup>

### 3.3 Empirische analyse van de uitbreidingsvraagprognoses naar sector

In deze paragraaf wordt de doeltreffendheid van de prognoses met betrekking tot de uitbreidingsvraag per bedrijfssector empirisch getoetst. Uitgangspunt zijn de bedrijfssectorprognoses van het CPB. Vervolgens zijn de procentuele ontwikkelingen en de prognoses van het aantal werkenden weergegeven ten opzicht van het aantal werkenden in het basisjaar 2002 volgens de EBB-data van het CBS. De EBB is een steekproefonderzoek. Zoals in ieder steekproefonderzoek hebben de uitkomsten een onnauwkeurigheidsmarge. De steekproefvariantie van het aantal werkzamen personen in elke beroep-sector-combinatie in de EBB is nogal hoog. Dit is vooral het geval voor beroep-sector-combinaties waar maar weinig personen werkzaam zijn. Het CBS schat regelmatig de 95%-betrouwbaarheidsmarges. In 1990 waren de geschatte betrouwbaarheidsintervallen<sup>9</sup> 30,1% voor 5.000 werkenden en 12,3% voor 30.000 werkenden. In 2003 liggen deze intervallen bij 31,3% en 12,8% respectievelijk, terwijl dit voor 2009 respectievelijk 27,7% en 11,3% is (CBS, 2010). Omdat de kwaliteit van deze sectorprognoses in hoge mate de kwaliteit van de prognoses per beroepsgroep (en per opleidingstype) kan bepalen, worden in deze paragraaf de bedrijfssectorprognoses onder de loep genomen.

---

8. De veronderstelde werkgelegenheidsontwikkeling voor 2003 en 2004 is gebaseerd op het *Centraal Economisch Plan 2003* van het CPB, voor de daaropvolgende jaren tot 2008 is gebruik gemaakt van de middellangetermijnramingen van het CPB.

9. De betrouwbaarheidsmarges hebben betrekking op de opgehoogde aantallen.

## De voorspelfouten

Tabel 3.1 geeft per bedrijfssector een overzicht van de werkgelegenheid in 2002, de prognoses voor 2007 en de realisatie voor 2007, de voorspelfout en het verlies per bedrijfssector. Uit de tabel blijkt dat de voorspelfout voor een aantal bedrijfssectoren negatief is. Dit betekent dat de prognoses voor deze beroepsgroepen de uitbreidingsvraag overschatten. Dit is voornamelijk het geval voor de sectoren *Handel en reparatie*, *Overheid en onderwijs*, *Overige industrie*, en *Bank- en verzekeringswezen*. Voor eerstgenoemde en laatstgenoemde sectoren werd zelfs een lichte stijging van de werkgelegenheid voorspeld, terwijl de werkgelegenheid in werkelijkheid licht is afgenomen.

**Tabel 3.1**

Verlies uitbreidingsvraagprognoses per bedrijfssector

Bedrijfssector	Referentie	Prognose	Realisatie	Voorspelfout	Verlies
	2002	2007	2007		
Landbouw en visserij	225.700	174.300	214.100	39.700	0,0310
Voeding	138.800	142.000	132.700	-9.300	0,0044
Chemie	107.600	104.300	128.800	24.600	0,0521
Metaal en elektrotechniek	380.100	355.600	395.600	40.000	0,0111
Overige industrie	357.300	339.400	292.200	-47.200	0,0175
Energie	52.800	57.600	58.300	800	0,0002
Bouw en onroerend goed	544.300	523.800	569.900	46.100	0,0072
Handel en reparatie	1.051.100	1.062.200	945.600	-116.600	0,0123
Transport en communicatie	411.600	394.700	447.900	53.100	0,0167
Bank- en verzekeringswezen	254.400	286.600	251.500	-35.100	0,0190
Horeca en zakelijke dienstverlening	1.177.600	1.228.100	1.216.900	-11.200	0,0001
Kwartaire diensten	1.161.300	1.266.200	1.299.200	33.000	0,0008
Overheid en onderwijs	975.700	1.089.000	1.010.100	-78.900	0,0065
Bedrijfssectoren totaal	7.035.000	7.224.400	7.259.400	34.900	0,0000

Voor een aantal bedrijfssectoren is de voorspelfout juist positief, wat betekent dat de groei in deze bedrijfssectoren is onderschat of de krimp is overschat. Het gaat hierbij voornamelijk om de sectoren *Transport en communicatie*, *Bouw en onroerend goed*, *Metaal en elektrotechniek*, en *Landbouw en visserij*. Voor de drie eerst-

genoemde sectoren werd zelfs lichte daling voorspeld, terwijl de werkgelegenheid in werkelijkheid licht is toegenomen. Afgezien van het totaal van bedrijfssectoren, is voor geen enkele afzonderlijke sector is de voorspelling opvallend goed (een verlies van 0,0000). De voorspellingen voor de sectoren *Energie* en *Horeca en zakelijke dienstverlening* zijn vrij redelijk (het verlies is kleiner dan 0,001).

### *De voorspelkwaliteit*

In Tabel 3.2 worden de bedrijfssectorprognoses vergeleken met de SAB-prognoses. Het gemiddeld verlies van de prognoses is groter dan het gemiddeld verlies van de SAB-prognose (zowel absoluut als relatief), wat leidt tot scores die groter zijn dan 1. Het verschil tussen het gemiddeld verlies van de absolute prognose en de relatieve prognose is zeer klein. Echter, aangezien het gemiddeld verlies van de relatieve prognose iets groter is dan die van de absolute prognose terwijl tegelijkertijd het gemiddeld verlies van de relatieve SAB-prognose iets kleiner is dan die van de absolute prognose, is de score van de relatieve prognose (1,19) iets hoger dan die van de absolute prognose (1,15). De voorspelkwaliteit van de sectorprognoses zijn dus slechter (zowel absoluut als relatief) dan die van de SAB-prognose.

Uit Tabel 3.2 blijkt verder dat zowel de absolute als relatieve sectorprognoses voor 1998 beter waren dan de prognoses voor 2007. Het verlies van de absolute prognoses is licht gestegen ten opzicht van de eerdere prognoses, terwijl het verlies van de relatieve prognoses sterk is gestegen ten opzicht van de eerdere prognoses. Opvallend is dat het verlies van de absolute SAB-prognoses is gehalveerd terwijl tegelijkertijd het verlies van de relatieve SAB-prognoses bijna is verdubbeld in vergelijking met eerdere prognoses. Dit betekent dat ondanks de verbetering van de absolute SAB-prognoses, de onderschatting van de totale werkgelegenheid er voor



zorgt dat de relatieve SAB-prognoses (die gecorrigeerd zijn voor het totale volume-effect) gestegen zijn.

**Tabel 3.2**

Totaaloverzicht voorspelkwaliteit uitbreidingsvraagprognoses per bedrijfssector

		Gem. verlies prognose	Gem. verlies SAB	Score
Prognoses	2007			
<i>Absoluut</i>		0,0085	0,0074	1,15
<i>Relatief</i>		0,0086	0,0073	1,19
Prognoses	1998			
<i>Absoluut</i>		0,0077	0,0148	0,52
<i>Relatief</i>		0,0020	0,0038	0,52
Prognoses	1994			
<i>Absoluut</i>		0,0036	0,0080	0,44
<i>Relatief</i>		0,0031	0,0072	0,44

### Verklaring van voorspelfouten

Tabel 3.3 geeft een verklaring voor de voorspelfout en de standaardafwijking van deze voorspelfout van de relatieve sectorprognoses.<sup>10</sup> Het blijkt dat de onderschattingscoëfficiënt niet significant van 0 verschilt. Er is dus geen sprake geweest van systematische over- of onderschatting. De schaal heeft wel een significant effect op de voorspelkwaliteit. Naarmate de omvang van een sector toeneemt zijn de sectorprognoses minder nauwkeurig. De standaardafwijking van de voorspelfout neemt iets meer dan proportioneel toe met de omvang van een bedrijfssector. Bij eerdere sectorprognoses werd eenzelfde effect van de schaal gevonden (Smits en Diephuis, 2001).

10. De voorspelfouten zijn gemeten door de logaritme te nemen van de absolute waarde van het verschil tussen de realisatie van 2007 en de prognose van 2007. De schaal is gemeten door de logaritme te nemen van de realisatie van 2007, terwijl de onderschattingscoëfficiënt gemeten is door het verschil te nemen van de prognose van 2007 en de waarde van de grootte in het basisjaar 2002 (de referentieprognose).

**Tabel 3.3**

Verklaring van voorspelfouten van de uitbreidingsvraagprognoses per sector (relatief)

Variabele	Parameter	T-waarde
Constante	-4,331	-1,02
Schaal	1,132	3,44 ***
Onderschattingscoëfficiënt	-0,000	-1,50

\*\*\* significant bij betrouwbaarheid van 99%

Naast bovengenoemde verklaringen is er echter nog een additionele factor die de voorspelfouten heeft beïnvloed. Een belangrijke factor is dat het CBS het aantal werkenden per beroepsgroep in 2002, gebaseerd op de Enquête Beroepsbevolking (EBB), heeft aangepast (deze aanpassing vond plaats nadat de prognoses waren opgesteld). Om de effecten van de veranderingen in de data ook in kaart te brengen voor bedrijfssectoren, is een regressie geschat waarin de voorspelfouten in de bedrijfssectorprognoses worden verklaard door de afwijking tussen de EBB-gegevens van het jaar 2002 die beschikbaar was ten tijde van het samenstellen van de prognoses in 2003 en de EBB-gegevens van het jaar 2002 die zijn gebruikt voor deze evaluatiestudie (gegevens uit 2009). De regressieresultaten zijn weergegeven in Tabel 3.4. De tabel laat zien dat de afwijking per sector tussen de EBB-data die voor het basisjaar beschikbaar waren ten tijde van de prognose (gegevens uit 2003) en na de bijstelling (gegevens uit 2009) de voorspelfout niet significant beïnvloedt. Hierbij moet worden opgemerkt dat er maar weinig sectoren zijn, wat sneller leidt tot niet-significante resultaten.

**Tabel 3.4**

Afwijking EBB als verklaring van voorspelfouten in de uitbreidingsvraag per sector (relatief)

Variabele	Parameter	T-waarde
Constante	3,042	0,64
Afwijking EBB-data	1,518	1,41

Naast de gegevens van het verlies van de uitbreidingsvraagprognoses per bedrijfssector in Tabel 3.1 zijn ook nog bedrijfssectorgegevens van het CPB<sup>11</sup> opgenomen in Bijlage 1 (Tabel B.1). Dit zijn de originele onbewerkte middellange-termijnramingen (MLT) en de realisaties zoals die nu in de Nationale Rekeningen (NR) staan. Uit de tabel is duidelijk te zien dat het verlies op basis van de MLT-ramingen en de realisaties in de NR ongeveer half zo hoog is als het verlies dat berekend is door ROA. Eén van de redenen hiervoor zou kunnen zijn dat de Nationale Rekeningen (NR)-gegevens moeten worden omgerekend naar EBB\_aantallen. Daarnaast zijn voor het basisjaar 2002 de herziene EBB-gegevens uit 2009 gebruikt zijn, terwijl voor de prognose van 2007 de voorspelde gegevens uit 2003 gebruikt zijn.

Daarom zijn naast de gegevens van het gemiddeld verlies van de uitbreidingsvraagprognoses en de SAB per bedrijfssector en de daarbij behorende score in Tabel 3.2, het gemiddeld verlies van de prognoses en de SAB en de bijbehorende score ook berekend door een correctie uit te voeren op de prognoses en de SAB (zie Tabel 3.5). De correctie op de prognoses bestaat uit het verschil tussen de EBB-gegevens uit 2003 en de herziene EBB-gegevens uit 2009 voor het basisjaar 2002. De gecorrigeerde prognose is gelijk aan de prognose minus de gegevens van 2002 uit 2003 plus de gegevens van 2002 uit 2009. Op deze manier is de prognoses beter te vergelijken met de realisatie die immers ook gebaseerd is op de recentere gegevens.

Tevens zijn aan de hand van een gecorrigeerde SAB ook het gemiddeld verlies van de SAB en de bijbehorende score berekend. De correctie op de SAB houdt in dat in plaats van het aantal werkenden in 2002 uitgegaan wordt van een trend<sup>12</sup>, wat een realistischere naïeve voorspelling is dan door simpel te veronderstellen dat de situatie in de prognosejaren gelijk is aan die van het basisjaar 2002. Bij het bepalen

---

11. Met dank aan Bert Smid van het CPB.

12. Voor meer informatie over de berekening van de trend, zie paragraaf 3.2 van Cörvers et al. (2004).

van de trend is uitgegaan van de ontwikkelingen in de jaren 1998-2002. Vervolgens is deze trend doorgetrokken voor de prognosejaren 2003 tot en met 2007. Uit Tabel 3.5 blijkt dat als de gecorrigeerde prognoses (EBB correctie) gebruikt worden, het gemiddeld verlies zowel in absolute als in relatieve termen een stuk lager wordt, wat tot lagere scores leidt (1,01 ten opzichte van respectievelijk 1,15 en 1,19). De gecorrigeerde SAB (trend in plaats van waarde basisjaar) leidt tot een lager absoluut maar een hoger relatief gemiddeld verlies. Als zowel de prognoses als de SAB gecorrigeerd worden leidt dit tot lagere scores (1,10 en 0,97 ten opzichte van 1,15 en 1,19) en dus een wat betere voorspelkwaliteit.

**Tabel 3.5**

Totaaloverzicht voorspelkwaliteit uitbreidingsvraagprognoses per bedrijfssector

		Gem. verlies prognose	Gem. verlies SAB	Score
Prognoses (zie ook Tabel 3.2)	2007			
<i>Absoluut</i>		0,0085	0,0074	1,15
<i>Relatief</i>		0,0086	0,0073	1,19
Prognoses gecorrigeerd	2007			
<i>Absoluut</i>		0.0075	0,0074	1,01
<i>Relatief</i>		0.0073	0,0073	1,01
SAB gecorrigeerd	2007			
<i>Absoluut</i>		0,0085	0.0068	1,25
<i>Relatief</i>		0,0086	0.0076	1,14
Prognoses en SAB gecorrigeerd	2007			
<i>Absoluut</i>		0.0075	0.0068	1,10
<i>Relatief</i>		0.0073	0.0076	0,97

Een andere optie om de effecten van veranderingen in EBB-data toch in kaart te brengen is om de EBB-data die voor het basisjaar beschikbaar waren ten tijde van de prognose (gegevens uit 2003) te gebruiken (in plaats van de gegevens uit 2009) om het gemiddeld verlies van de prognoses en de SAB, alsook de score uit te

rekenen. Deze resultaten, op dezelfde wijze berekend als in Tabel 3.2, zijn te vinden in Tabel B.2 van Bijlage 1. Uit deze resultaten blijkt dat als de EBB-data uit 2003 gebruikt wordt de score in absolute termen verbetert (score van 1,01) en in relatieve termen verslechtert (score van 1,38) ten opzichte van de gegevens in Tabel 3.2 waarin de recentere EBB-data uit 2009 gebruikt wordt. Hieruit blijkt dat het veel verschil uit kan maken welke EBB-gegevens gebruikt worden in de analyses, en of er wel of niet gecorrigeerd wordt voor veranderingen in EBB-data.

### *Conclusies*

Uit de evaluatie van de uitbreidingsvraagprognoses per sector kan de conclusie worden getrokken dat de kwaliteit van de uitbreidingsvraagprognoses over het algemeen achteruit is gegaan. De SAB-prognose blijkt zowel absoluut als relatief gezien een betere voorspeller te zijn van de uitbreidingsvraagprognoses per sector dan de werkelijke prognose. Tevens blijkt de schaal een significant positief effect te hebben op de voorspelfouten. Een pluspunt is dat de relatieve prognoses van de uitbreidingsvraag per sector niet duiden op een systematische over- of onderschatting. Ondanks dat de afwijking per sector van de EBB-data geen significant effect laat zien op de voorspelfouten (zie Tabel 3.4), blijkt dat als de gecorrigeerde prognoses en SAB gehanteerd worden, het gemiddeld verlies en bijbehorende score een stuk lager liggen (zie Tabel 3.5). Hieruit blijkt dat het veel verschil uit kan maken welke EBB-gegevens gebruikt worden in de analyses en of er wel of niet gecorrigeerd wordt voor veranderingen in EBB-data.

### 3.4 Empirische analyse van de uitbreidingsvraagprognoses naar beroep

In deze paragraaf wordt de doeltreffendheid van de prognoses met betrekking tot de uitbreidingsvraag per beroepsgroep empirisch getoetst.

## *De voorspelfouten*

Tabel 3.6 geeft een overzicht van de prognoseresultaten met betrekking tot de uitbreidingsvraag per beroepsgroep. In de tabel staan achtereenvolgens het aantal werkzame personen per beroepsgroep in 2002 (de referentie), de prognose en de realisatie van het aantal werkzame personen in 2007, alsmede de voorspelfout en het verlies per beroepsgroep.

Voor veel beroepsgroepen is de voorspelfout negatief. Dit betekent dat de prognoses voor deze beroepsgroepen de gerealiseerde uitbreidingsvraag overschatten. In meer detail, voor 56% van de beroepsgroepen wordt een overschatting gevonden. De vijf beroepsgroepen met de grootste nominale overschatting van de uitbreidingsvraag zijn: *Boekhouders en secretaresses*, *Managers*, *Systeemanalisten*, *Programmeurs* en *Leidinggevenden*. De uitbreidingsvraag werd voor deze beroepsgroepen met minimaal 22.500 overschat. De prognoses voor de genoemde groepen tezamen resulteren in een voorspelfout van meer dan 154.000, wat 38% van de totale overschatting vormt.

Tegenover deze overschattingen staan enkele beroepsgroepen waarvoor de uitbreidingsvraagprognoses van 2007 een onderschatting waren. De voornaamste uitschieters, in absolute aantallen, zijn hier: *Verkopers*, *Verplegenden en doktersassistenten*, *Economen*, *Kunstenaars*, en *Medewerkers sociaal-cultureel werk en personeel en arbeid*. Voor de *Verkopers* wordt zelfs een onderschatting van meer dan 62.000 waargenomen. Voor de andere beroepsgroepen ligt de overschatting tussen de 22.400 en de 16.500. In totaal wordt er voor 44% van de beroepsgroepen een onderschatting waargenomen. Daarnaast blijkt uit de tabel dat er voor één beroepsgroep, namelijk *Medisch analisten*, een zeer goede voorspelling is gedaan: er is bijna geen verlies (0.000).

Voor sommige beroepsgroepen wordt de werkgelegenheidstoename van 2002 tot en met 2007 slecht voorspeld door het model. De beroepsgroep *Economen* heeft een verlies van 4,3482. In plaats van de voorspelde daling van de werkgelegenheid met 900 werkenden vond er een stijging plaats van de werkgelegenheid met bijna 19.600 werkenden. De werkgelegenheid in deze beroepsgroep is erg gevoelig voor de conjunctuur terwijl het model alleen de effecten van de structurele economische variabelen kan meenemen. Cyclische variatie vertroebelt het effect van structurele economische variabelen en leiden eerder tot slechte werkgelegenheidsvoorspellingen van het model. Andere beroepsgroepen die problematisch zijn: *Politie-inspecteurs en officieren*, *Elektrotechnici*, *Docent 1<sup>e</sup>-graads zonder specialisatie*, en *Organisatiekundigen*. Een belangrijke factor die de grote verliezen van *Docent 1<sup>e</sup>-graads zonder specialisatie* en *Politie-inspecteurs en officieren* grotendeels bepaald is het kleine aantal werkenden in deze beroepsgroepen. De slechte voorspelling van *Economen* werd eveneens waargenomen in het voorafgaande evaluatierapport van de uitbreidingsvraag (zie Dupuy, 2005), alhoewel het verlies toen niet zo groot was (0,338) en de richting van de voorspelfout tegenovergesteld was. Dat wil zeggen dat er toen in plaats van de voorspelde stijging van de werkgelegenheid met 600 werkenden een daling plaatsvond van 16.000 naar 10.000 werkenden.

Naast de gegevens in Tabel 3.6 is er een correlatie uitgerekend en getoetst om te zien of er een relatie is tussen de richting van de prognoses (de voorspelde stijging of daling in de werkgelegenheid) aan de ene kant en de gerealiseerde richting (de daadwerkelijke stijging of daling in de werkgelegenheid) aan de andere kant. Dit levert een correlatie op van 0,39 met een p-waarde van 0,0000. Dit betekent dat bij een betrouwbaarheidsinterval van 99% er inderdaad een statistisch significante relatie is tussen de richting van de prognoses en de gerealiseerde richting.

**Tabel 3.6**

Verlies uitbreidingsvraagprognoses per beroepsgroep

<b>Beroepsgroep</b>	<b>Referentie</b>	<b>Prognose</b>	<b>Realisatie</b>	<b>Voorspelfout</b>	<b>Verlies</b>
	<b>2002</b>	<b>2007</b>	<b>2007</b>		
Leraar basisonderwijs	144.100	164.000	173.800	9.800	0,0046
Docenten exacte, medische en verzorgende vakken (2e en 3e gr.)	22.700	25.900	19.800	-6.100	0,0717
Docenten exacte, medische en verzorgende vakken (1e gr. en WO)	11.400	13.100	13.800	700	0,0034
Docenten landbouw en techniek (2e en 3e gr.)	14.600	17.400	13.900	-3.600	0,0595
Docenten landbouw en techniek (1e gr. en WO)	6.000	5.300	2.300	-3.000	0,2464
Docenten economisch-administratieve vakken (2e en 3e gr.)	12.800	15.900	9.900	-6.000	0,2188
Docenten economisch-administratieve vakken (1e gr. en WO)	7.100	9.500	5.800	-3.800	0,2826
Docenten talen en expressie	42.700	45.800	39.900	-5.900	0,0193
Docenten letteren (1e gr. en WO)	20.800	23.100	21.600	-1.500	0,0053
Docenten sociale vakken (2e en 3e gr.)	14.800	15.300	13.600	-1.700	0,0129
Docenten sociale vakken (1e gr. en WO)	6.100	7.300	5.900	-1.400	0,0505
Docenten 2e en 3e graads zonder specialisatie	9.300	9.500	9.700	200	0,0005
Docent 1e graads zonder specialisatie	2.800	3.400	1.700	-1.700	0,3819
Onderwijskundig medewerkers	9.600	11.400	10.600	-700	0,0061
Onderwijskundigen en pedagogen	21.700	25.200	27.500	2.400	0,0120
Rij-instructeurs	9.000	8.000	6.200	-1.800	0,0386
Sportinstructeurs	11.500	12.900	15.700	2.800	0,0598
Tolken, vertalers en schrijvers	12.300	13.200	10.100	-3.100	0,0629
Bibliotheekassistenten	21.000	20.200	15.000	-5.300	0,0628
Bibliothecarissen	5.900	6.200	7.200	1.000	0,0277
Grafisch ontwerpers	11.400	12.900	15.800	2.900	0,0636
Kunstenaars	54.200	51.900	70.100	18.200	0,1125



<b>Beroepsgroep</b>	<b>Referentie</b>	<b>Prognose</b>	<b>Realisatie</b>	<b>Voorspelfout</b>	<b>Verlies</b>
	<b>2002</b>	<b>2007</b>	<b>2007</b>		
Geestelijk verzorgers	1.500	1.400	1.500	100	0,0061
Geestelijken	3.900	3.900	2.700	-1.200	0,0928
Journalisten	20.900	22.600	22.300	-300	0,0001
Taalkundigen	8.000	8.400	7.500	-900	0,0125
Agrarische hulparbeiders	5.700	5.500	6.100	600	0,0119
Agrarische arbeiders	101.700	88.900	98.400	9.500	0,0088
Agrarische vakkrachten	8.800	7.500	7.200	-300	0,0015
Milieuhygiënisten en agrarisch vertegenwoordigers	12.800	13.300	12.500	-800	0,0040
Landbouwmachinebestuurders en vissers	11.700	9.000	8.100	-900	0,0053
Agrarische bedrijfshoofden	121.600	91.700	105.900	14.200	0,0136
Productiemedewerkers	92.400	84.900	99.100	14.200	0,0236
Laboratorium-assistenten	2.200	2.400	2.800	300	0,0242
Laboranten	12.900	13.400	11.400	-2.000	0,0240
Technisch analisten	6.400	7.400	10.500	3.100	0,2302
Natuurwetenschappers	14.700	16.400	9.300	-7.100	0,2295
Conciërges	26.000	28.600	22.900	-5.700	0,0478
Hoofden technische dienst	13.700	13.500	9.200	-4.300	0,1004
Werktuigbouwkundigen	8.400	8.700	10.100	1.400	0,0279
Bouwvakkers	188.100	185.500	179.200	-6.400	0,0011
Aannemers en installateurs	198.700	202.700	189.700	-13.000	0,0043
Architecten en bouwkundig projectleiders	44.200	50.600	56.300	5.600	0,0162
Weg- en waterbouwkundigen	20.400	22.900	24.100	1.200	0,0034
Weg- en waterbouwkundige arbeiders	24.400	21.600	21.100	-400	0,0003
Weg- en waterbouwkundige vakkrachten	34.400	34.200	28.600	-5.600	0,0261

<b>Beroepsgroep</b>	<b>Referentie 2002</b>	<b>Prognose 2007</b>	<b>Realisatie 2007</b>	<b>Voorspelfout</b>	<b>Verlies</b>
Weg- en waterbouwkundig ontwerpers en projectleiders	11.800	13.000	9.900	-3.100	0,0667
Metaalarbeiders	77.000	72.900	70.000	-2.900	0,0014
Bankwerkers en lassers	64.800	64.500	62.700	-1.800	0,0008
Bedrijfshoofden metaalbewerking	4.700	4.200	2.400	-1.800	0,1543
Assembleurs	24.400	22.900	15.100	-7.800	0,1019
Monteurs	212.400	197.800	202.500	4.700	0,0005
Werktuigbouwkundig ontwerpers en hoofden technische dienst	26.300	24.200	25.400	1.200	0,0022
Elektronicamonteurs	8.300	9.300	6.800	-2.500	0,0921
Monteurs en controleurs elektrotechnische producten	14.700	15.400	11.800	-3.600	0,0599
Elektromonteurs	88.600	85.300	80.400	-4.900	0,0031
Elektrotechnisch ontwerpers en bedrijfshoofden	12.100	12.200	10.500	-1.700	0,0201
Elektrotechnici	8.700	9.500	3.600	-5.900	0,4660
Grafisch productiepersoneel	21.300	21.300	15.600	-5.700	0,0720
Grafische vakkrachten	30.300	24.800	25.800	1.000	0,0011
Mechanisch operators	63.500	58.000	66.300	8.200	0,0168
Procesoperators	47.300	49.300	50.300	1.000	0,0004
Procestechnologen	9.800	9.200	7.900	-1.200	0,0158
Materiaalkundigen	11.400	10.700	9.500	-1.200	0,0104
Confectie-arbeiders	29.900	26.300	19.700	-6.600	0,0487
Schoen- en kleermakers	6.000	5.400	3.200	-2.100	0,1304
Laders en lossers	123.800	135.200	142.400	7.200	0,0034
Chauffeurs	245.400	236.000	234.800	-1.200	0,0000
Schippers en conducteurs	15.200	12.300	15.400	3.000	0,0399
Vliegers, scheepskapiteins en leidinggevendende transport	17.900	17.200	16.000	-1.200	0,0047

<b>Beroepsgroep</b>	<b>Referentie 2002</b>	<b>Prognose 2007</b>	<b>Realisatie 2007</b>	<b>Voorspelfout</b>	<b>Verlies</b>
Stewards	9.100	10.700	9.100	-1.600	0,0312
Verpleeghulpen en leerling-verpleegkundigen	30.900	33.100	26.600	-6.500	0,0439
Verplegenden en doktersassistenten	112.800	112.800	135.200	22.400	0,0394
Therapeuten en verpleegkundigen	115.700	123.400	128.900	5.400	0,0022
Artsen	68.000	75.800	72.900	-2.800	0,0017
Apothekersassistenten en medisch laboranten	47.500	51.000	47.600	-3.400	0,0052
Medisch analisten	18.200	21.600	21.600	0	0,0000
Apothekers	7.200	7.800	7.100	-700	0,0105
Afdelingshoofden zorginstelling	7.900	8.500	11.700	3.300	0,1746
Kantoorhulpen, inpakkers en colporteurs	58.400	56.600	69.300	12.700	0,0471
Ondersteunende administratieve hulpkrachten	10.400	9.100	11.200	2.100	0,0404
Bedrijfshoofden	51.400	49.200	64.000	14.800	0,0826
Economen	9.800	8.900	29.400	20.500	4,3482
Productieplanners	56.300	69.100	72.400	3.300	0,0035
Organisatie-adviseurs	50.500	58.900	65.900	7.000	0,0190
Organisatiedeskundigen	22.600	25.800	38.700	12.900	0,3288
Receptionisten en administratieve employés	247.100	248.200	236.300	-12.000	0,0023
Boekhouders en secretaresses	458.300	508.700	468.600	-40.100	0,0077
Assistent accountants	104.200	101.300	98.600	-2.700	0,0007
Accountants	32.700	34.100	46.700	12.700	0,1498
Verzekeringsagenten	17.700	18.000	11.800	-6.300	0,1257
Commercieel employés	293.100	280.700	296.400	15.700	0,0029
Commercieel medewerkers	234.700	232.900	240.300	7.400	0,0010
Technisch-commercieel employés	15.500	14.600	15.700	1.200	0,0057

<b>Beroepsgroep</b>	<b>Referentie</b>	<b>Prognose</b>	<b>Realisatie</b>	<b>Voorspelfout</b>	<b>Verlies</b>
	<b>2002</b>	<b>2007</b>	<b>2007</b>		
Technisch-bedrijfskundig medewerkers	11.900	11.700	15.700	4.000	0,1123
Juridisch en fiscaal medewerkers	36.700	41.300	34.400	-6.900	0,0352
Juridisch, bestuurlijk medewerkers	19.600	20.800	10.100	-10.700	0,2994
Juristen	66.800	72.100	76.700	4.600	0,0047
Administratieve transportemployés	27.400	28.100	23.900	-4.100	0,0227
Leidinggevend	55.100	62.200	39.700	-22.500	0,1660
Managers	106.600	119.300	82.700	-36.600	0,1178
Medisch secretaresses	24.400	24.800	17.400	-7.400	0,0928
Programmeurs	85.800	103.100	76.200	-26.800	0,0979
Systeemanalisten	138.700	166.200	137.600	-28.500	0,0424
Informatici	38.100	43.300	34.000	-9.300	0,0598
Technisch systeemanalisten	13.700	11.800	11.200	-600	0,0018
Activiteitenbegeleiders en medewerkers arbeidsbemiddeling	76.000	83.600	98.600	14.900	0,0386
Medewerkers sociaal-cultureel werk en personeel en arbeid	92.000	101.600	118.100	16.500	0,0322
Hoofden sociaal-cultureel werk en personeel en arbeid	9.700	11.100	14.600	3.600	0,1352
Sociaal-wetenschappelijk medewerkers	13.700	14.800	21.500	6.700	0,2415
Sociaal-wetenschappelijk onderzoekers	30.000	34.600	42.900	8.300	0,0774
Vakkenvullers	39.000	40.400	31.000	-9.500	0,0587
Interieurverzorgers	179.100	184.000	190.100	6.100	0,0012
Verkopers	265.100	273.800	336.100	62.300	0,0553
Winkeliers	146.900	123.100	116.200	-6.900	0,0022
Hulpkrachten horeca en verzorging	176.300	185.000	191.300	6.300	0,0013
Ziekenverzorgenden	79.500	96.000	99.600	3.700	0,0021
Verzorgend personeel	211.600	218.300	234.600	16.300	0,0059

<b>Beroepsgroep</b>	<b>Referentie</b>	<b>Prognose</b>	<b>Realisatie</b>	<b>Voorspelfout</b>	<b>Verlies</b>
	<b>2002</b>	<b>2007</b>	<b>2007</b>		
Café- en snackbarhouders	18.900	13.100	12.100	-1.000	0,0029
Bedrijfshoofden horeca	55.200	57.000	52.000	-5.100	0,0084
Bakkers en slaggers	13.700	12.700	10.400	-2.200	0,0268
Aspirant politieagenten, soldaten en beveiligingshulpkrachten	48.000	52.900	52.800	-100	0,0000
Politieagenten, onderofficieren en beveiligingsemployés	50.800	52.900	60.000	7.100	0,0198
Politie-inspecteurs en officieren	5.000	5.500	10.800	5.300	1,1439
Brandweerlieden	7.500	6.400	8.500	2.100	0,0800
Beroepsgroepen totaal	7.035.000	7.229.400	7.259.400	30.000	0,0000

Tabel 3.7 geeft een overzicht van de grootste voorspelfouten van de uitbreidingsvraagprognoses per beroepsgroep. Uit de gegevens blijkt dat de beroepsgroepen met de grootste overschatting (negatieve voorspelfout) of onderschatting (positieve voorspelfout) niet systematisch terug keren in de “top 5”. Alleen voor *Boekhouders en secretaresses* werd zowel in deze evaluatie als in de vorige evaluatie (Dupuy, 2005) een vrij grote overschatting gevonden. Voor de andere beroepsgroepen in de “top 5” van grootste overschattingen werd in de vorige evaluatie juist een onderschatting gevonden. Ook voor de “top 5” van onderschattingen kan geconcludeerd worden dat deze beroepsgroepen niet systematisch onderschat worden.

**Tabel 3.7**

Grootste voorspelfouten uitbreidingsvraagprognoses per beroepsgroep

Beroepsgroep	Voorspelfout	Voorspelfout vorige evaluatie
Boekhouders en secretaresses	-40.100	-24.200
Managers	-36.600	8.800
Systeemanalisten	-28.500	36.200
Programmeurs	-26.800	22.300
Leidinggevenden	-22.500	12.600
Verkopers	62.300	-14.500
Verplegenden en doktersassistenten	22.400	100
Economen	20.500	-6.100
Kunstenaars	18.200	-600
Medewerkers sociaal-cultureel werk en personeel en arbeid	16.500	9.500

### *De voorspelkwaliteit*

Tabel 3.8 geeft een totaaloverzicht van de voorspelkwaliteit van de prognoses van de uitbreidingsvraag van 2007. De tabel laat zien dat het gemiddelde verlies van de absolute prognoses is toegenomen ten opzichte van het gemiddeld verlies van de absolute prognoses tot 2002. Het gemiddelde verlies van de absolute prognoses is echter nog steeds lager dan die van de prognoses tot 2000. Aangezien het absolute gemiddelde verlies van de SAB voorspelling minder gestegen is dan het absolute gemiddelde verlies van de prognose, is de score van het absolute gemiddelde verlies gestegen. Het gemiddeld verlies van de relatieve prognoses is flink gedaald ten

opzichte van het gemiddeld verlies van de relatieve prognoses tot 2002. Voor de relatieve prognoses is er een verbetering opgetreden. Aangezien het relatieve gemiddelde verlies van de SAB minder gedaald is dan het relatieve gemiddelde verlies van de prognose, is de score van het relatieve gemiddelde verlies gedaald.

**Tabel 3.8**

Totaaloverzicht voorspelkwaliteit uitbreidingsvraagprognoses per beroepsgroep

		Gem. verlies prognose	Gem. verlies SAB	Score
Prognoses	2007			
<i>Absoluut</i>		0,0293	0,0304	0,96
<i>Relatief</i>		0,0295	0,0314	0,94
Prognoses	2002			
<i>Absoluut</i>		0,0268	0,0300	0,89
<i>Relatief</i>		0,0399	0,0412	0,97
Prognoses	2000			
<i>Absoluut</i>		0,0309	0,0430	0,72
<i>Relatief</i>		0,0214	0,0280	0,77

De scores van de prognoses geven aan dat de werkelijke prognoses voor de absolute en de relatieve uitbreidingsvraag van 2007 beter waren dan de SAB (scores van 0,96 en 0,94). De score van de absolute prognoses van 2007 is verslechterd ten opzichte van de score van de absolute prognoses tot 2002. De score van de relatieve prognoses is verbeterd ten opzichte van de prognoses tot 2002, maar is niet zo goed als die van tot 2000. Opvallend is dat de absolute prognoses een slechtere score geven dan de relatieve prognoses, een resultaat dat sinds 1994 niet meer is waargenomen. Het feit dat de score van de absolute prognoses slechter is dan die van de relatieve prognoses duidt er op dat bij de prognoses een groter deel van het verlies veroorzaakt wordt door de onderschatting van de totale werkgelegenheid dan bij de SAB. De relatieve score corrigeert als het ware voor de algemene conjunctuur, oftewel voor de algemene tendens om de werkgelegenheid te onderschatten, vandaar dat de relatieve score (0,94) beter is dan de absolute score (0,96).

### *Kwaliteit afzonderlijke modelstappen*

Zoals in paragraaf 3.2 is beschreven, worden de prognoses per beroepsgroep in twee stappen gemaakt. Eerst worden de prognoses per bedrijfssector vertaald naar prognoses per beroepssegment. In de tweede stap worden de prognoses per beroepssegment vervolgens verbijzonderd naar de onderliggende beroepsgroepen. Het is interessant om na te gaan in welke mate beide stappen bijdragen tot een verbetering van de prognoses.

Allereerst wordt gekeken wat de prognoseresultaten zouden zijn geweest als de beroepenstructuur constant zou zijn verondersteld. De voorspelde groei van een beroep binnen een bedrijfssector is dan gelijk aan de totale voorspelde groei van de bedrijfssector. Dat betekent dat verondersteld wordt dat de werkgelegenheidsontwikkelingen per beroep volledig worden bepaald door verschuivingen tussen bedrijfssectoren. Vervolgens wordt gekeken wat de resultaten zouden zijn geweest als wel rekening zou zijn gehouden met veranderingen in de beroepenstructuur van een bedrijfssector maar niet met verschillen tussen beroepsgroepen binnen een beroepssegment. De groei per beroepsgroep is dan gelijk aan de totale groei van het beroepssegment waar het onder valt.

Tabel 3.9 geeft de voorspelkwaliteit van de verschillende stappen in het model. De eerste rij geeft de SAB-prognose. In dat geval wordt zowel de beroepenstructuur als de bedrijfssectoropbouw in de economie constant verondersteld. Vervolgens vergelijken we de kwaliteit van deze prognose met de prognose waarbij wel rekening is gehouden met werkgelegenheidsverschuivingen tussen bedrijfssectoren maar niet met veranderingen in de beroepenstructuur binnen een bedrijfssector. Het gemiddeld verlies stijgt sterk. Deze prognose bevat 104% van de fout van de SAB. Oftewel, de SAB voorspelt beter. Dit komt overeen met eerdere resultaten met betrekking tot de sectorprognoses (zie bijvoorbeeld Tabel 3.2). Als vervolgens rekening wordt gehouden met veranderingen in de beroepenstructuur wordt een kleine verbetering behaald. Het gemiddeld verlies bevat nu nog 98% van het verlies



van de voorgaande stap. In de volgende stap, waarbij ook veranderingen binnen een beroepssegment een rol spelen verbetert de voorspelkwaliteit nog verder (94%).

**Tabel 3.9**

Betekenis van de verschillende modelstappen voor de uitbreidingsvraagprognoses naar beroepsgroep (absoluut)

	Gemiddeld verlies	Score t.o.v. voorgaande stap
SAB	0,0304	
Beroepenstructuur 2002	0,0317	1,04
Segmentstructuur 2002	0,0312	0,98
Prognose	0,0293	0,94

Het is duidelijk dat de grootste winst in voorspelkwaliteit wordt behaald door rekening te houden met verschuivingen binnen de segmentstructuur. Het reeds gebruikte random-coëfficiëntenmodel dat gebruikt wordt om de gegevens op segmentniveau te verbijzonderen naar beroepsgroep zorgt met een score van 0,94 voor de beste voorspelkwaliteit. Bij de prognoses van 2002 bleek juist dat de beste voorspelkwaliteit werd behaald op basis van de bedrijfssectorprognoses. De verbetering in de voorspelkwaliteit was toen nihil als rekening werd gehouden met verschuivingen binnen beroepenstructuur en segmentstructuur (zie Dupuy, 2005). Het nieuwe model van Cörvers en Dupuy (2003) biedt een betere micro-economische onderbouwing en maakt beter gebruik van sectorgegevens (kapitaal, R&D e.d.) dan het oude model. Het model wordt gebruikt om de sectorgegevens te verbijzonderen naar beroepssegment. De score van 0,98 lijkt er op te wijzen dat het nieuwe model ook daadwerkelijk heeft bijgedragen aan de verbetering van de voorspelkwaliteit, alhoewel de verbetering minimaal is ten opzichte van de vorige evaluatie (Dupuy, 2005) die een score van 0,99 opleverde.

## *Verklaring van voorspelfouten*

Tabel 3.10 geeft een verklaring voor de voorspelfout en de standaardafwijking van deze voorspelfout van de relatieve uitbreidingsvraagprognoses per beroepsgroep.<sup>13</sup> De onderschattingscoëfficiënt is significant van nul en heeft een negatief effect. Het negatieve effect geeft aan dat er bij de uitbreidingsvraagprognoses per beroepsgroep sprake is van een significante overschatting van afwijkingen van de referentieprognose. Deze systematische overschatting van de veranderingen impliceert waarschijnlijk dat de prognoses relatief veel ruis bevatten. Bij eerdere voorspellingen (voor 1998) werd ook een systematische overschatting gevonden. Sinds 1998 wordt gebruik gemaakt van het random-coëfficiëntenmodel. Omdat dit model bij het extrapoleren van trends rekening houdt met de statistische betrouwbaarheid ervan, bevatten de prognoses minder ruis. De resultaten in Tabel 3.8 laten nu voor de eerste keer sinds het gebruik van het random-coëfficiëntenmodel een significant effect van overschatting zien, alhoewel de grootte van de parameter erg laag is (-0,000).

De parameter voor het schaaleffect is positief en kleiner dan één, hetgeen impliceert dat grote beroepsgroepen in aantallen slechter, maar in procenten beter voorspeld worden dan de kleine beroepsgroepen. Wanneer deze resultaten vergeleken worden met de voorgaande evaluatie (zie Dupuy, 2005), dan blijkt het schaaleffect te zijn toegenomen (van 0,502 naar 0,989). De coëfficiënt van de uitwijkmogelijkheden van beroepsgroepen over bedrijfssectoren is negatief maar niet significant. Hetzelfde geldt voor de substitutiemogelijkheden van opleidingen en de conjunctuurgevoeligheid.

---

13. De voorspelfouten zijn gemeten door de logaritme te nemen van de absolute waarde van het verschil tussen de realisatie van 2007 en de prognose van 2007. De schaal is gemeten door de logaritme te nemen van de realisatie van 2007, terwijl de onderschattingscoëfficiënt gemeten is door het verschil te nemen van de prognose van 2007 en de waarde van de grootte in het basisjaar 2002 (de referentieprognose).

**Tabel 3.10**

Verklaring van voorspelfouten van de uitbreidingsvraagprognoses per beroepsgroep (relatief)

Variabele	Parameter	T-waarde
Constante	0,156	2,83 ***
Schaal	0,989	200,73 ***
Substitutiemogelijkheden opleidingen	-0,002	-1,56
Uitwijkmogelijkheden naar bedrijfssectoren	-0,003	-1,59
Conjunctuurgevoeligheid	-0,008	-0,59
Onderschattingscoëfficiënt	-0,000	-7,67 ***

\*\*\* significant bij betrouwbaarheid van 99%

Naast bovengenoemde verklaringen is er echter nog een additionele factor die de voorspelfouten heeft beïnvloed. Een belangrijke factor is dat het CBS het aantal werkenden per beroepsgroep in 2002, gebaseerd op de Enquête Beroepsbevolking (EBB), heeft aangepast (deze aanpassing vond plaats nadat de prognoses waren opgesteld). Om de effecten van de veranderingen in de data ook in kaart te brengen voor alle beroepsgroepen, is een regressie geschat waarin de voorspelfouten in de uitbreidingsvraag per beroepsgroep worden verklaard door de afwijking tussen de EBB-gegevens van het jaar 2002 die beschikbaar was ten tijde van het samenstellen van de prognoses in 2003 en de EBB-gegevens van het jaar 2002 die zijn gebruikt voor deze evaluatiestudie (gegevens uit 2009). De regressieresultaten zijn weergegeven in Tabel 3.11. De tabel laat zien dat de afwijking per beroepsgroep tussen de EBB data die voor het basisjaar beschikbaar waren ten tijde van de prognose en na de bijstelling de voorspelfout significant beïnvloedt. Een toename van de afwijking in de EBB-data met 1% leidt tot een toename van de voorspelfout met 1,012%.

**Tabel 3.11**

Afwijking EBB als verklaring van voorspelfouten in de uitbreidingsvraag per beroepsgroep (relatief)

Variabele	Parameter	T-waarde
Constante	6,777	11,60 ***
Afwijking EBB-data	1,012	5,87 ***

\*\*\* significant bij betrouwbaarheid van 99%

## Kwalitatieve typeringen

Bij de uitbreidingsvraag per beroepsgroep (UBG) is een kwalitatieve typering gemaakt. De indeling van de kwalitatieve typeringen is als volgt (Cörvers et al, 2004):

min.	≤	UBG	≤	10%	erg laag
10%	<	UBG	≤	30%	laag
30%	<	UBG	≤	70%	gemiddeld
70%	<	UBG	≤	90%	hoog
90%	<	UBG	≤	100%	erg hoog

Dit leidt tot de volgende grenswaarden (UV is de uitbreidingsvraag en de grenswaarden stellen de uitbreidingsvraag als percentage van de werkgelegenheid per beroepsgroep voor):

min.	≤	UV	≤	-2,2	erg laag
-2,2	<	UV	≤	-0,7	laag
-0,7	<	UV	≤	1,1	gemiddeld
1,1	<	UV	≤	1,9	hoog
1,9	<	UV	≤	max.	erg hoog

**Tabel 3.12**

Typeringen van de uitbreidingsvraag per beroepsgroep (relatief)

<i>Prognose</i>	<i>Realisatie</i>					Totaal
	Erg laag	Laag	Gemiddeld	Hoog	Erg hoog	
Erg laag	5	3	0	1	2	11
Laag	11	5	5	0	1	22
Gemiddeld	21	6	9	3	14	53
Hoog	6	2	4	1	12	25
Erg hoog	4	2	1	0	8	15
Totaal	47	18	19	5	37	126

Tabel 3.12 toont in hoeverre de kwalitatieve typering van de relatieve voorspelling overeenkomt met de gerealiseerde typering. Uit de tabel blijkt dat voor 28 van de 126 beroepsgroepen (22%) de voorspelde typering overeenkomt met de realisatie, terwijl voor 44 beroepsgroepen geldt dat de voorspelde kwalitatieve typering één categorie te hoog of te laag is. Dit betekent dat er in totaal in 57% van de beroepsgroepen een geheel of bijna juiste typering gemaakt is. Dit percentage is lager dan het percentage van de uitbreidingsvraagprognoses voor 2002 die in 61% van de gevallen geheel of bijna juist bleken te zijn. Er is met name een sterke toename van beroepsgroepen die in de realisatie respectievelijk een “erg laag” of juist “erg hoog” uitbreidingsvraag bleken te hebben maar die in de prognose de typering “gemiddeld” kregen (dit komt respectievelijk 21 keer en 14 keer voor). Het lijkt er op dat de werkelijke waarden van de uitbreidingsvraag als percentage van de werkgelegenheid per beroepsgroep een grotere variantie bevatten dan de voorspelde waarden.

Daarnaast kan gekeken worden in welke mate de kwalitatieve typering van de relatieve voorspelling helemaal niet overeenkomt met de gerealiseerde typering. Voor 6 beroepsgroepen (5%) geldt dat de voorspelde kwalitatieve typering maar liefs vier categorieën te hoog of te laag is. Voor 10 beroepsgroepen (8%) geldt dat de voorspelde kwalitatieve typering drie categorieën te hoog of te laag is. Dit betekent dat er in totaal in 13% van de beroepsgroepen een geheel onjuiste typering gemaakt is. Dit is een verbetering ten opzichte van de 15% van de vorige evaluatie (Dupuy, 2005).

Naast de gegevens in Tabel 3.10 is er ook een correlatie uitgerekend en getoetst om te zien of er een relatie is tussen de voorspelde typering en de werkelijke typering. Dit levert een correlatie op van 0,29 met een t-waarde van 3,4. Dit betekent dat bij een betrouwbaarheidsinterval van 99% er inderdaad een statistisch significante relatie is tussen de richting van de prognoses en de gerealiseerde richting. In het vorige evaluatierapport werd dezelfde relatie gevonden (de correlatie was 0,31 met een t-waarde van 3,7).

## Conclusies

In de arbeidsmarktprognoses die in 2003 zijn samengesteld is voor het eerst gebruik gemaakt van een nieuw model voor de uitbreidingsvraag naar beroep.<sup>14</sup> In dit nieuwe model wordt de beroepenstructuur binnen bedrijfssectoren geschat aan de hand van, behalve de tijdtrend, vier factoren, te weten: het productieniveau ('non-homothetic production function') per bedrijfssector, de kapitaalintensiteit per bedrijfssector ('capital-skill complementarity'), het gebruik van nieuwe meer kennisintensieve technologieën ('skill biased technological change') en de relatieve lonen op de verschillende arbeidsmarktsegmenten. Door de uitbreidingsvraagprognoses die destijds gemaakt zijn geworden te evalueren kan gekeken worden in hoeverre het nieuwe model ook goed voorspelt.

Uit de evaluatie van de uitbreidingsvraagprognoses per beroepsgroep kan de conclusie getrokken worden dat de kwaliteit van de absolute prognoses is verslechterd terwijl die van de relatieve prognoses is verbeterd. Als naar de voorspelkwaliteit van de afzonderlijke modelstappen wordt gekeken blijkt dat er winst in voorspelkwaliteit wordt behaald door rekening te houden met verschuivingen in de segmentstructuur (0,98) en, in meerdere mate, door rekening te houden met verschuivingen tussen beroepsgroepen binnen de segmentstructuur (0,94). Ondanks dat het reeds gebruikte random-coëfficiënten model (score van 0,94) voor de grootste winst in voorspelkwaliteit zorgt, draagt het nieuwe model van Cörvers en Dupuy (2003, 2010) met een score van 0,98 ook bij aan de verbetering van de voorspelkwaliteit, alhoewel de verbetering minimaal is ten opzichte van de vorige evaluatie (Dupuy, 2005) die een score van 0,99 opleverde.

In tegenstelling tot voorafgaande evaluaties duidt de relatieve prognose van de uitbreidingsvraag per beroepsgroep op systematische overschatting van veranderingen, alhoewel de grootte van de parameter erg laag is (-0,000). Net als

---

14. Zie Cörvers et al. (2004),; Cörvers en Dupuy (2006). Het destijds ontwikkelde model is overigens nog verder ontwikkeld door rekening te houden met interacties tussen beroep-sector-combinaties: Cörvers en Dupuy (2010).

voorafgaande evaluaties is het schaaleffect significant positief echter kleiner dan één, hetgeen impliceert dat grote beroepsgroepen in aantallen slechter, maar in procenten beter voorspeld worden dan de kleine beroepsgroepen. Het schaaleffect blijkt wel te zijn toegenomen. Uitwijkmogelijkheden van beroepsgroepen over bedrijfssectoren, substitutiemogelijkheden van opleidingen en de conjunctuurgevoeligheid spelen geen significante rol. Verder blijkt dat afwijkingen in de EBB-data een significant positief effect hebben op voorspelfouten.

Met betrekking tot de kwalitatieve typering van de relatieve voorspelling van de uitbreidingsvraag per beroepsgroep kan gezegd worden dat er in totaal in 57% van de beroepsgroepen een geheel of bijna juiste typering gemaakt is, wat lager is dan in de voorafgaande evaluatie (61%) (Dupuy, 2005). Waarschijnlijk komt dit door het feit dat de werkelijke waarden van de uitbreidingsvraag als percentage van de werkgelegenheid per beroepsgroep een grotere variantie bevatten dan de voorspelde waarden. Als gekeken wordt in welke mate de kwalitatieve typering van de relatieve voorspelling helemaal niet overeenkomt met de gerealiseerde typering blijkt dat dit is een verbetering oplevert (13%) ten opzichte van de 15% van de vorige evaluatie (Dupuy, 2005).

## **4 De indicator toekomstige arbeidsmarktsituatie (ITA) tot 2008**

### **4.1 Inleiding**

De arbeidsmarktperspectieven voor een opleiding zijn zowel afhankelijk van vraagfactoren als van aanbodfactoren. De verhouding tussen de verwachte vraag- en aanbodstromen komt tot uitdrukking in de Indicator Toekomstige Arbeidsmarktsituatie (ITA). Deze indicator geeft de verwachte arbeidsmarktperspectieven per opleidingstype weer. De indicator heeft een belangrijke verandering ondergaan sinds de prognoses van 1994. Voorheen werd alleen rekening gehouden met de ex ante vraag en aanbod voor een opleiding en niet met de (passieve) substitutievraag

als gevolg van tekorten of overschotten bij andere opleidingen. Sindsdien wordt deze substitutievraag wel meegenomen. In dit hoofdstuk wordt met behulp van de schoolverlaters-enquêtes BVE-Monitor, HBO-Monitor en WO-Monitor geëvalueerd of de voorspelde perspectieven daadwerkelijk tot uiting komen in de positie van de schoolverlaters op de arbeidsmarkt.

## 4.2 Prognosemethodiek

Voor een goed inzicht in het arbeidsmarktperspectief van een opleiding zijn twee zaken van belang. Ten eerste de huidige arbeidsmarktsituatie van de opleiding. Discrepanties die op een bepaald moment bestaan kunnen nog enige tijd doorwerken. Als er in het basisjaar sprake is van een aanbodoverschot dan kan dat er toe leiden dat een deel van de arbeidsmarktinstroom werkloos wordt of een baan buiten het eigen domein moet accepteren. Personen die geen passende baan hebben gevonden blijven wellicht toch nog enige tijd op zoek naar een baan in het eigen beroepsdomein. Ze concurreren daarom nog wel met schoolverlaters die later de arbeidsmarkt op stromen. De vraag is natuurlijk hoe lang deze situatie voortduurt. Er komt een bepaald moment waarop personen die werkloos zijn of werkzaam zijn in een beroep buiten het eigen domein door de werkgevers niet meer als geschikt aanbod worden beschouwd op dat beroepsdomein. Voor een aanbodtekort geldt een zelfde redenering. Als werkgevers een deel van de vraag niet op bevredigende wijze vervuld krijgen, dan zal dit in eerste instantie leiden tot openstaande vacatures. Als vacatures te lang open staan zullen ze er vaak toe overgaan om personeel met een andere opleidingsachtergrond aan te trekken of om het werk op een andere wijze te organiseren. Verwacht mag worden dat de huidige discrepanties vooral van invloed zijn op het arbeidsmarktperspectief op de korte termijn en minder van belang voor de middellange termijn.

Ten tweede is ook de verwachte verandering in de vraag/aanbodverhoudingen van belang voor het arbeidsmarktperspectief. Deze factor komt volledig tot uitdrukking in de ITA. Daarnaast is ook het effect van de huidige arbeidsmarktsituatie



gedeeltelijk meegenomen bij het bepalen van de ITA. Aan de aanbodzijde wordt behalve de voorspelde arbeidsmarktinstroom van schoolverlaters ook het aantal kortdurige werklozen met het betreffende opleidingstype aan het begin van de prognoseperiode meegeteld. Er is dus verondersteld dat personen die langer dan een jaar werkloos zijn geen serieuze concurrentie voor schoolverlaters meer vormen. Een tekort aan het begin van de prognose periode werkt daarentegen niet door in de ITA. Openstaande vacatures worden niet tot de vraag gerekend. Vandaar dat een ITA die tussen 1,00 en 1,05 in ligt als ‘redelijk’ wordt gezien. De ITA is als volgt gedefinieerd:

$$ITA = (E_t + INS_{t:t+5} + W_t) / (E_t + BO_{t:t+5} + SV_{t:t+5}) \quad (4.1)$$

Waarbij:

$E_t$  = aantal werkenden in  $t$

$INS_{t:t+5}$  = voorspelde arbeidsmarktinstroom voor de periode  $t:t+5$

$W_t$  = aantal kortdurig werklozen in  $t$

$BO_{t:t+5}$  = voorspelde baanopeningen voor de periode  $t:t+5$

$SV_{t:t+5}$  = voorspelde substitutievraag voor de periode  $t:t+5$

Aan de vraagzijde wordt naast de voorspelde uitbreidings- en vervangingsvraag ook de (passieve) substitutievraag (als gevolg van tekorten of overschotten bij andere opleidingen) meegerekend. De voorspelde baanopeningen bestaan uit de positieve uitbreidingsvraag en de vervangingsvraag.

### 4.3 Empirische analyse van de ITA

De ITA geeft een indicatie van het arbeidsmarktperspectief voor een opleidingstype. Zoals gezegd wordt bij het bepalen van de ITA het effect van de huidige arbeidsmarktpositie van een opleiding maar in beperkte mate meegenomen, er wordt alleen rekening gehouden met kortdurige werkloosheid. De ITA kan geïnterpreteerd worden als de ratio van het verwachte aanbod over 5 jaar ten opzichte van de verwachte vraag over 5 jaar. Een slechte arbeidsmarktpositie

betekent dat het verwachte aanbod de verwachte vraag zal overtreffen. In deze paragraaf zullen we eerst nagaan in welke mate de ITA een goede indicator is voor de verandering van de arbeidsmarktpositie tussen  $t$  en  $t+5$ .

Vraag- en aanboddyscrepanties op de arbeidsmarkt leiden maar ten dele tot werkloosheid of openstaande vacatures. Als er voor een bepaalde opleiding sprake is van een aanbodoverschot, dan zullen personen met deze opleidingsachtergrond genoeg nemen met minder gunstige arbeidsvoorwaarden. Ze zullen banen onder hun niveau en tegen een lagere beloning aanvaarden. Bij een tekort aan personen met een bepaalde opleidingsachtergrond zullen werkgevers betere arbeidsvoorwaarden gaan aanbieden om vacatures toch vervuld te krijgen. Daarnaast zullen ze ook personen rekruteren met een andere opleidingsachtergrond dan de gewenste en deze mensen door middel van bij- of omscholing alsnog geschikt maken voor de functie. In welke mate vraag -en aanbodoverschotten dus tot respectievelijk openstaande vacatures of werkloosheid leiden is afhankelijk van de flexibiliteit van de arbeidsmarkt en de uitwijk- en substitutiemogelijkheden. Bij een evaluatie van de voorspelde arbeidsmarktperspectieven is het van belang met dergelijke processen rekening te houden.

De eerste indicator voor de arbeidsmarktdyscrepanties die we analyseren is de ex post gap. Deze gap is gedefinieerd als het percentage schoolverlaters dat niet in het eigen domein werk heeft gevonden maar of werkloos is geworden of buiten het eigen domein werkzaam is.<sup>15</sup> De verhouding tussen de ex ante gap (de werkelijke discrepantie) en de ex post gap (de waargenomen discrepantie) is onder meer afhankelijk van de flexibiliteit van de arbeidsmarkt. Als het loon en andere arbeidsmarktvoorwaarden zich in het geheel niet aanpassen zal de ex post gap gelijk zijn aan de ex ante gap. Nu zal ook, indien er geen sprake is van een aanbodoverschot, toch een aantal schoolverlaters buiten hun eigen beroepsdomein

---

15. Of schoolverlaters in het eigen domein werkzaam zijn, is bepaald op basis van het door de werkgevers vereiste opleidingsniveau en -richting. Schoolverlaters die werkzaam zijn in een baan onder hun opleidingsniveau en/of buiten de eigen vakrichting, werken niet in hun eigen domein.

gaan werken of werkloos worden. Bijvoorbeeld omdat banen buiten hun domein toch hun voorkeur hebben of omdat ze achteraf minder geschikt bleken te zijn voor banen binnen het domein. Bij een tekort zal de gap daarom weliswaar kleiner worden maar hij zal nooit gelijk aan 0 worden omdat altijd een deel van de schoolverlaters sowieso werkloos wordt of buiten het eigen domein terechtkomt.<sup>16</sup>

Overigens valt te verwachten dat niet alle opleidingen op een zelfde wijze op arbeidsmarktdiscrepanties reageren. In sommige segmenten is de markt flexibeler of zijn er meer uitwijk- en/ of substitutiemogelijkheden dan in andere segmenten. Daar we alleen een voorspelling hebben van de discrepantie over de totale periode 2003-2008 en niet over de afzonderlijke jaren kunnen we alleen een vergelijking maken tussen jaren en niet tussen opleidingen. Er moet dus verondersteld worden dat alle opleidingen op eenzelfde manier op arbeidsmarktdiscrepanties reageren.

De schoolverlaters-enquêtes BVE-Monitor, HBO-Monitor en WO-Monitor geven informatie over de verschillende aspecten van de arbeidsmarktpositie van recent afgestudeerden. Deze databronnen zijn daarom bij uitstek geschikt voor een evaluatie van de ITA. Er is echter niet voor alle opleidingstypes informatie beschikbaar, waardoor de volgende analyses niet systematisch 104 opleidingstypen hebben.

Ten eerste is er gekeken of er een verband is tussen de verandering in de gap tussen 2002 en 2007 en de ITA en tussen het niveau van de gap in 2007 en de ITA. De volgende vergelijkingen zijn hiervoor geschat:

$$\Delta Gap_j = \alpha + \beta(ITA_j - 1) + e_j \quad (4.2)$$

$$Gap_{j,07} = \alpha' + \beta'(ITA_j - 1) + e'_j \quad (4.3)$$

---

16. Merk op dat de gap alleen betrekking heeft op een aanbodoverschot. Voor een analyse van het effect van een vraagoverschot zouden gegevens over de vraagzijde van de arbeidsmarkt geanalyseerd moeten worden, zoals vacature cijfers. Deze gegevens zijn echter niet op dit aggregatieniveau beschikbaar.

waarbij  $e_j$  een error term is,  $j$  het onderwijstype aangeeft en *Gap* het percentage schoolverlaters aangeeft dat niet in het eigen domein werk heeft gevonden.

Merk op dat gap alleen betrekking heeft op een aanbodoverschot. Voor een analyse van het effect van een vraagoverschot zouden gegevens over de vraagzijde van de arbeidsmarkt geanalyseerd moeten worden, zoals vacature cijfers. Deze gegevens zijn echter niet op dit aggregatieniveau beschikbaar. In de vergelijking wordt de ITA gerelateerd aan de situatie waarin vraag en aanbod in evenwicht zijn. Vergelijking (4.2) analyseert het verband tussen de verandering in arbeidsmarktpositie tussen 2002 en 2008. Als de ITA een waarde van 1 heeft dan is de verandering aan de aanbodzijde even groot als aan de vraagzijde. Een hogere waarde van de ITA betekent een verslechtering van de arbeidsmarktsituatie; het verwachte teken van  $\beta$  is daarom positief. De constante term  $\alpha$  geeft de algehele verandering in de gap onafhankelijk van vraag- en aanbodfactoren. In vergelijking (4.3) is het verwachte teken van  $\beta$  ook positief. Hier kan  $\alpha$  geïnterpreteerd worden als de natuurlijke waarde van de gap.

Tabel 4.1 geeft de schattingresultaten. Het blijkt dat voor zowel vergelijking (4.2) als vergelijking (4.3) beide coëfficiënten significant van 0 verschillen. In vergelijking (4.2) is het teken van  $\beta$  significant positief, er is dus een positief verband gevonden tussen het voorspelde arbeidsmarktperspectief en de verandering in de arbeidsmarktpositie tussen 2002 en 2008. Vergelijking (4.3) laat een significant positief effect zijn voor beide coëfficiënten. Dit betekent dat er een positief verband is gevonden tussen de ITA en de hoogte van de gap in 2007. De natuurlijke waarde van de gap in 2007 is 49,2. Dat betekent dat gemiddeld ruim 49% van de schoolverlaters werkloos is of buiten het eigen domein werkzaam is in 2007.

**Tabel 4.1**

Relatie tussen voorspelde arbeidsmarktpositie en huidige arbeidsmarktsituatie

Indicator arbeidsmarktsituatie	$\alpha$	T-waarde	$\beta$	T-waarde
$\Delta Gap$	-7,003	-3,45 ***	27,384	3,85 **
$Gap_{07}$	49,187	17,21 ***	48,656	3,05 ***

\*\* significant bij betrouwbaarheid van 95%

\*\*\* significant bij betrouwbaarheid van 99%

De ITA lijkt dus niet alleen een indicator te zijn voor de absolute arbeidsmarktpositie van een opleiding, maar ook voor de relatieve arbeidsmarktpositie van een opleiding. Dit in tegenstelling tot voorgaande evaluaties, die slechts significante resultaten lieten zien voor het niveau van de gap en niet voor de verandering in de gap.

Naast de gegevens in Tabel 4.1 is het verband tussen de verandering in de gap en de ITA en tussen het niveau van de gap en de ITA ook bekeken voor een ITA zonder het aantal kortdurig werklozen en een ITA zonder de voorspelde substitutievraag (zie Bijlage 1, Tabel B.3). De resultaten laten zien dat de variabelen qua niveau weliswaar niet veel veranderen, maar qua significantie wel. De ITA lijkt een minder significante indicator te zijn voor zowel de absolute als de relatieve arbeidsmarktpositie van een opleiding zodra het aantal kortdurig werklozen en de voorspelde substitutievraag uit de definitie van de ITA gehaald worden.

Op basis van de schoolverlaterenquêtes (BVE-Monitor, HBO-Monitor en WO-Monitor), wordt voor verschillende aspecten van de arbeidsmarktpositie de relatie met de ITA geschat. De volgende arbeidsmarktaspecten zijn daarbij in beschouwing genomen: het werkloosheidspercentage, het lange termijn werkloosheidspercentage (percentage schoolverlaters dat langer dan 4 maanden werkloos was), het percentage schoolverlaters dat onder het eigen opleidingsniveau werkzaam is, het percentage schoolverlaters dat buiten de eigen vakrichting werkzaam is, het percentage schoolverlaters met vast werk, het percentage

schoolverlaters met deeltijdwerk en het gemiddeld (bruto) maandelijks loon (in 1.000 euro's). Per aspect is de volgende vergelijking geschat:

$$y_j^k = \alpha^k + \beta^k (ITA_j - 1) + e_j^k \quad (4.4)$$

waarbij  $y_j^k$  = arbeidsmarktaspect  $k$  voor opleidingstype  $j$ .

Daar ook in bovenstaande vergelijkingen de ITA wordt gerelateerd aan de evenwichtssituatie, kan  $\alpha^k$  geïnterpreteerd worden als de 'natuurlijke' omvang van de betreffende indicator. De parameter  $\beta^k$  geeft aan hoe de indicator wijzigt als gevolg van de veranderende perspectieven.

**Tabel 4.2**

Relatie tussen voorspelde arbeidsmarktpositie en huidige arbeidsmarktsituatie

Indicator arbeidsmarktsituatie	$\alpha$	T-waarde			$\beta$	T-waarde	
<i>Werkloosheid</i>	0,948	44,90	***		0,021	4,84	***
<i>Lange termijn werkloosheid</i>	0,999	34,91	***		0,002	0,70	
<i>Onderbenutting</i>	0,985	33,23	***		0,001	1,31	
<i>Buiten eigen richting</i>	0,912	32,70	***		0,004	4,42	***
<i>Deeltijdwerk</i>	0,953	32,34	***		0,002	2,78	***
<i>Vast werk</i>	1,253	19,46	***		-0,004	-3,79	***
<i>Gemiddeld loon</i>	1,285	27,17	***		-0,000	-5,99	***

\*\*\* significant bij betrouwbaarheid van 99%

Tabel 4.2 geeft de uitkomsten van de afzonderlijke schattingen.<sup>17</sup> Het blijkt wederom dat er sprake is van een significant positieve relatie tussen het werkloosheidspercentage en de ITA. Alhoewel lange termijn werkloosheid (meer dan 4 maanden werkloos na afstuderen) niet significant schijnt te worden beïnvloed door het arbeidsmarktperspectief. Vijf van de zeven indicatoren blijken een significant effect te ondervinden van de voorspelde perspectieven (de ITA): het percentage schoolverlaters dat werkloos is, het percentage schoolverlaters dat buiten de eigen

17. Aangezien de regressoren voor alle vergelijkingen gelijk zijn, zijn de hier gepresenteerde OLS resultaten equivalent aan de schattingen van simultane verklaringsmodellen voortkomend uit Seemingly Unrelated Regression (SUR).

richting werkzaam is, het percentage schoolverlaters met deeltijdwerk, het percentage schoolverlaters met vast werk en het gemiddeld loon van schoolverlaters.

Voor de eerste drie indicatoren is er sprake van een significant positieve relatie met het voorspelde arbeidsmarktperspectief. Dit betekent dat een hoge ITA, oftewel een slecht arbeidsmarktperspectief, leidt tot een hoger percentage schoolverlaters dat werkloos is, buiten de eigen richting werkzaam is en in deeltijd werkt. Voor de laatste twee indicatoren is er sprake van een significant negatieve relatie met het voorspelde arbeidsmarktperspectief. Het gemiddeld loon laat een significant negatief verband zien. Een hoge ITA leidt tot een daling in het gemiddeld loon van schoolverlaters. Bovenstaande resultaten werden ook in voorgaande evaluaties gevonden. Zoals in voorgaande evaluatie studies is er een negatief verband tussen de ITA en het percentage schoolverlaters met vast werk. Echter, in tegenstelling tot vorige evaluaties, berust deze relatie op een significante schatting. Een hoge ITA leidt tot een daling in het percentage schoolverlaters met vast werk.

Naast de gegevens in Tabel 4.2 is de relatie tussen de voorspelde arbeidsmarktpositie en de huidige arbeidsmarktsituatie ook bekeken voor een ITA zonder het aantal kortdurig werklozen en een ITA zonder de voorspelde substitutievraag (zie Bijlage 1, Tabel B.4). De resultaten laten zien dat de variabelen qua niveau niet veel veranderen. Qua significantie verandert er met name iets voor het percentage schoolverlaters met deeltijdwerk, deze indicator voor de arbeidsmarktsituatie is veel minder significant voor een ITA zonder kortdurig werklozen en/of een ITA zonder substitutievraag. Interessant is om te zien dat voor een ITA zonder substitutievraag het lange termijn werkloosheidspercentage (percentage schoolverlaters dat langer dan 4 maanden werkloos was) significant wordt, alhoewel dit slechts is bij een betrouwbaarheidsinterval van 90%. Door rekening te houden met de substitutievraag (als gevolg van tekorten of overschotten bij andere opleidingen) is het effect van de ITA op het lange termijn werkloosheidspercentage niet meer significant.

Om te bepalen hoe goed de ITA de toekomstige arbeidsmarktsituatie heeft voorspeld moeten alle aspecten van deze arbeidsmarktsituatie simultaan worden beschouwd. We willen de volgende vergelijking schatten:

$$ITA_{j,07} = \alpha_{07} + \sum \beta^k_{07} y^k_{j,07} + e_{j,07} \quad (4.5)$$

waarbij  $y^k_{j,07}$  = arbeidsmarktaspect  $k$  voor opleidingstype  $j$  in 2007.

Dit is dus eigenlijk een omgekeerde aanpak vergeleken met vergelijking (4.4). Het gaat hier dan ook niet om een causaal verband, maar om de mate waarin afzonderlijke indicatoren informatie geven over de meest aannemelijke waarde van de niet direct waargenomen ‘realisatie’ van de ITA. Daarom is een regressievergelijking geschat waarin de ITA wordt verklaard uit bovengenoemde indicatoren van de arbeidsmarktsituatie.<sup>18</sup> Op basis van de schattingen van vergelijking (4.5) zijn de realisaties van de ITA bepaald. De schatting geeft ook een score-indicatie, omdat de  $\bar{R}^2$  (de aangepaste R-kwadraat) van deze schatting kan worden vergeleken met de score-verhouding volgens de vergelijking: score =  $1 - \bar{R}^2$ . Omdat de  $\bar{R}^2 = 0,34$  is, zou de op deze wijze bekende score 0,66 zijn. De ITA scores voor de perioden 1995-2000, 1997-2002, 1999-2004 en 2001-2006 waren respectievelijk gelijk aan: 0,80, 0,86, 0,98 en 0,96. De score stijgt dus over de tijd, wat een indicatie is voor een verslechtering van de voorspelkwaliteit van de ITA. Echter, Dupuy (2009) argumenteert dat de stijgende trend kunstmatig is en geen indicatie is voor een verslechtering van de voorspelling maar de moeilijkheid benadrukt van het meten van de kwaliteit van de ITA-voorspelling (zie voor meer informatie Dupuy, 2009). De ITA score van 2007 (de periode van 2002-2008) laat nu voor de eerste keer sinds 2000 een daling zien. Dit is een indicatie voor een verbetering van de voorspelkwaliteit van de ITA.

---

18. Coëfficiënten voor 2007 zijn: *constante* = 1,306, werkloosheid = 0,015, lange termijn werkloosheid = -0,004, onderbenutting = -0,007, buiten eigen richting = 0,009, deeltijdwerk = 0,005, vast werk = -0,003 en gemiddeld loon = -0,000.



Vergelijking (4.5) wijst er op dat ongeveer 39% (de onaangepaste  $R^2 = 0,39$ ) van de standaardafwijking in de ITA over opleidingstypen samen valt met de standaardafwijking in de niveaus van de diverse arbeidsmarktaspecten  $k$  over de opleidingstypen  $j$ . Dit betekent dat 61% van de standaardafwijking in the ITA over opleidingstypen niet samen valt met de standaarddeviatie in deze arbeidsmarkt-aspecten. Inderdaad, de ITA-voorspellingen zijn afgeleid van een structureel model en zijn afhankelijk van voorspellingen van de uitbreidingsvraag, vervangingsvraag en aanbod per opleidingstypen. Het structurele model houdt rekening met complexe structurele veranderingen in de arbeidsmarkt zoals technische veranderingen (in de uitbreidingsvraag). De arbeidsmarktaspecten houden niet direct rekening met deze veranderingen. Vergelijking (4.5) is in feite een (partiële) reduced form vergelijking. Dit betekent dat de beta coëfficiënten in vergelijking (4.5) niet gebruikt kunnen worden om de structurele parameters van het model te identificeren, aangezien ze zelf combinaties van de structurele parameters en/ of variabelen van het model zijn. Dit punt wordt geïllustreerd door het feit dat die coëfficiënten veranderen over de tijd als de structuur van het model beïnvloed wordt door veranderingen in de economie. Bijvoorbeeld, technologische verandering beïnvloed de uitbreidingsvraag op complexe manieren die leiden tot substitutie tussen opleidingen van werknemers en misschien eventueel er toe leiden dat sommige werknemers onder hun opleidingsniveau of buiten hun eigen specialisatie-richting gaan werken of zelfs werkloos raken.

Om het belang te laten zien van deze structurele veranderingen in de ITA voorspelling, worden de veranderingen in de ITA tussen twee voorspelperioden, zeg 2007 en 2005, uitgesplitst in structurele veranderingen en veranderingen in de arbeidsmarktaspecten. Dit wordt gedaan door vergelijking (4.5) te schatten met behulp van een zogenaamd “pooled regression model”, waarbij de jaren 2002, 2005 en 2007 gezamenlijk in één model geanalyseerd worden. De afzonderlijke jaren kunnen dan als dummyvariabelen worden opgenomen.

Door toe te laten dat de coëfficiënten kunnen variëren van jaar tot jaar, is het mogelijk om een aantal t-toetsen uit te voeren op de  $\beta$ -coëfficiënten (per jaar) alsook een simpele F-toets voor de stabiliteit van de parameters. Als vergelijking (4.5) een structureel verband weergeeft dan is de verwachting dat de  $\beta$ -coëfficiënten constant zijn over de tijd, oftewel de veranderingen in de ITA tussen twee perioden slechts bepaald worden door veranderingen in arbeidsmarktaspecten  $y^k$ .

**Tabel 4.3**

Pooled regression van de ITA op de arbeidsmarktaspecten

Jaar	2002		2005		2007	
Indicator arbeidsmarktsituatie	$\beta$		$\beta_3$		$\beta_5$	
	Reference					
Constante	1,306	***	-0,261	*	-0,264	**
<i>Werkloos</i>	0,015	***	-0,014	**	0,000	
<i>&gt; 4 maanden werkloos</i>	-0,004		0,005		-0,003	
<i>Onderbenutting</i>	-0,001		-0,001		0,000	
<i>Buiten eigen richting</i>	0,001		0,002		0,001	
<i>Deeltijdwerk</i>	0,000		-0,000		0,001	
<i>Vast werk</i>	-0,003	**	0,003	***	0,00	
<i>Gemiddeld loon</i>	0,000	***	-0,000		0,00	
$R^2$	0,40					
$R^2_{adj}$	0,35					
$N$	298					
$F$ -test	3,01	$p = 0,000$				

\* significant bij betrouwbaarheid van 90%

\*\* significant bij betrouwbaarheid van 95%

\*\*\* significant bij betrouwbaarheid van 99%

De nul-hypothese van constante parameters over de tijd kan verworpen worden bij een niveau van 0% zoals aangegeven wordt door de F-stat in Tabel 4.3. Als we kijken naar de standvastigheid van iedere parameter afzonderlijk, dan laten de t-toetsen ook zien dat er significante variatie over de tijd is voor de constante, het werkloosheidspercentage en het percentage schoolverlaters met vast werk. Deze analyse wijst er op dat ondanks dat de ITA voor jaar  $t$  significant gecorreleerd is met de waargenomen arbeidsmarktsituatie in jaar  $t$  (voor elk jaar  $t$  wordt 40% van de variatie in ITA tussen opleidingstypen verklaard door variatie in arbeidsmarkt-aspecten zoals werkloosheid, vast werk, etc. zoals aangegeven door de  $R^2$  in Tabel 4.3), de voorspelling niet alleen rekening houdt met veranderingen in werk-

loosheid, vast werk, etc. maar ook met meer complexe structurele veranderingen. Deze bevindingen werden ook gevonden in voorafgaande evaluaties (Dupuy, 2009).

### *Kwalitatieve typering*

Op basis van de geschatte realisatie van de ITA zoals weergegeven in vergelijking (4.5), kunnen we net als in paragraaf 3.4 een vergelijking maken tussen de typering van de prognoses en de realisaties. Bij de ITA per opleidingstype is een kwalitatieve typering gemaakt. De indeling van de kwalitatieve typering is als volgt (Cörvers et al, 2004):

min.	≤	ITA	≤	0,85	zeer goed
0,85	<	ITA	≤	1,00	goed
1,00	<	ITA	≤	1,05	redelijk
1,05	<	ITA	≤	1,15	matig
1,15	<	ITA	≤	max.	slecht

Tabel 4.4 geeft deze vergelijking van de kwalitatieve typering van het arbeidsmarktperspectief voor de 104 opleidingen waarvoor schoolverlatersgegevens beschikbaar zijn. Hierbij moet echter wel bedacht worden dat de vergelijking die het gewicht van de diverse indicatoren van de arbeidsmarktsituatie bepaalt, geschat is op basis van de prognoses, waarbij dus is aangenomen dat de voorspelfouten een verwachting van 0 hebben. Beide delen, niet alleen de prognoses maar ook de geschatte realisaties, bevatten een zekere mate van onzekerheid. De eventuele onzuiverheden in de typering zijn daarom niet alleen toe te schrijven aan fouten in de prognoses, discrepanties tussen prognose en realisatie, als gevolg van het schattingsresultaat, kunnen ook de oorzaak zijn van een onnauwkeurige bepaling van de gerealiseerde perspectieven.

De tabel toont in hoeverre de kwalitatieve typering van de voorspelling overeenkomt met de gerealiseerde typering per opleidingstype. Uit de tabel blijkt dat voor 45 van de 104 opleidingstypen (43%) de voorspelde typering overeenkomt met de realisatie. Dit percentage is lager dan het percentage van de ITA-prognoses voor 2002 (Dupuy, 2005) dat in 64% van de gevallen geheel juist bleek te zijn, maar hoger dan het percentage van de vorige ITA-prognose voor 2004 (Dupuy, 2009) dat in 41% van de gevallen geheel juist bleek te zijn. Voor 39 opleidingstypen geldt dat de voorspelde kwalitatieve typering één categorie te hoog of te laag is. Dit betekent dat er in totaal in 2007 in 81% van de opleidingstypen een geheel of bijna juiste typering, wat een verbetering is ten opzicht van de 77% in de vorige evaluatie (Dupuy, 2009).

**Tabel 4.4**  
Kwalitatieve typering van de ITA per onderwijstype

<i>Prognose</i>	<i>Realisatie</i>					Totaal
	Zeer goed	Goed	Redelijk	Matig	Slecht	
Zeer goed	0	5	3	0	0	8
Goed	1	24	10	5	6	46
Redelijk	0	2	1	5	1	9
Matig	0	2	9	11	4	26
Slecht	0	1	2	3	9	15
Totaal	1	34	25	24	20	104

Naast de gegevens in Tabel 4.4 is er ook een correlatie uitgerekend en getoetst om te zien of er een relatie is tussen de voorspelde typering en de werkelijke typering. Dit levert een correlatie op van 0,53 met een t-waarde van 6,2. Dit betekent dat bij een betrouwbaarheidsinterval van 99% er inderdaad een statistisch significante relatie is tussen de richting van de prognoses en de gerealiseerde richting. In het vorige evaluatierapport was er nauwelijks correlatie gevonden (de correlatie was 0,06 met een t-waarde van 0,6).

Naast de gegevens in Tabel 4.4 is de vergelijking van de kwalitatieve typering van het arbeidsmarktperspectief ook bekeken voor een ITA zonder het aantal kortdurig werklozen en een ITA zonder de voorspelde (passieve) substitutievraag (zie Bijlage

1, tabel B.5). De resultaten laten zien dat voor de ITA 43% de voorspelde typering overeenkomt met de realisatie, terwijl voor 38% opleidingstypen geldt dat de voorspelde kwalitatieve typering één categorie te hoog of te laag is (in totaal 81%). Voor de ITA zonder kortdurig werklozen is dit respectievelijk 41% en 37% (in totaal 78%) en voor de ITA zonder substitutievraag is dit respectievelijk 30% en 47% (in totaal 77%). Uit deze resultaten blijkt dus dat het uitmaakt of de substitutievraag wordt meegenomen of niet: als de substitutievraag wordt meegenomen stijgt het aantal juist getypeerde opleiding van 30% naar 43%, waarbij het totaal aantal opleidingstypen met een geheel of bijna juiste typering stijgt van 77% naar 81%. Het maakt dus veel verschil uit of de substitutievraag wel of niet wordt meegenomen in de ITA. Bij de evaluatie van 2001 bleek dat door het opnemen van de (passieve) substitutievraag de kwaliteit van de voorspelde arbeidsmarktperspectieven licht verbeterde, namelijk 22 in plaats van 19 juiste typeringen. Echter als één typering er naast ook meegenomen werd was het aantal goede typeringen gelijk, namelijk 35 (Smits en Diephuis, 2001). In tegenstelling tot de evaluatie van 2001 zien we nu dus dat niet alleen het aantal juiste typeringen toeneemt door de substitutievraag mee te nemen, maar dat ook het aantal goede typeringen (juiste plus één er naast) stijgt.

### *Conclusies*

De ITA indicatoren zijn de kern van de voorspellingen per opleiding en beroep. In de evaluatie van de ITA is er een positief verband gevonden tussen het voorspelde arbeidsmarktperspectief en de verandering in de arbeidsmarktpositie tussen 2002 en 2008. Ook is er een positief verband gevonden tussen de ITA en de hoogte van de gap in 2007. De natuurlijke waarde van de gap in 2007 is 49,2. Dat betekent dat gemiddeld ruim 49% van de schoolverlaters buiten het eigen domein werkzaam is in 2007. De ITA lijkt dus niet alleen een indicator te zijn voor de absolute arbeidsmarktpositie van een opleiding, maar ook voor de relatieve arbeidsmarktpositie van een opleiding. Dit in tegenstelling tot voorgaande evaluaties, die slechts significante resultaten lieten zien voor het niveau van de gap en niet voor de verandering in de gap.

Voor de afzonderlijke schattingen blijkt wederom dat er sprake is van een significant positieve relatie tussen het werkloosheidspercentage en de ITA. Alhoewel lange termijn werkloosheid (meer dan 4 maanden werkloos na afstuderen) niet significant schijnt te worden beïnvloed door het arbeidsmarktperspectief. Vijf van de zeven indicatoren blijken een significant effect te ondervinden van de voorspelde perspectieven (de ITA): het percentage schoolverlaters dat werkloos is, het percentage schoolverlaters dat buiten de eigen richting werkzaam is, het percentage schoolverlaters met deeltijdwerk, het percentage schoolverlaters met vast werk en het gemiddeld loon van schoolverlaters.

Om te bepalen hoe goed de ITA de toekomstige arbeidsmarktsituatie heeft voorspeld zijn alle aspecten van de arbeidsmarktsituatie simultaan geschat. De ITA score van 2007 (de periode van 2002-2008) laat nu voor de eerste keer sinds 2000 een daling zien (0,66 t.o.v. 0,96 voor 2001-2006). Dit is een indicatie voor een verbetering van de voorspelkwaliteit van de ITA.

Daarnaast blijkt dat veranderingen in de ITA tussen twee perioden niet slechts bepaald worden door veranderingen in arbeidsmarktaspecten  $y^k$  (zoals weergegeven door de *F-stat*). De t-toetsen laten ook zien dat er significante variatie over de tijd is voor de constante, het werkloosheidspercentage en het percentage schoolverlaters met vast werk. Deze analyse wijst er op dat ondanks dat de ITA voor jaar  $t$  significant gecorreleerd is met de waargenomen arbeidsmarktsituatie in jaar  $t$  de voorspelling niet alleen rekening houdt met veranderingen in werkloosheid, vast werk, etc. maar ook met meer complexe structurele veranderingen. Deze bevindingen werden ook gevonden in voorafgaande evaluaties (Dupuy, 2009).

Met betrekking tot de kwalitatieve typering van de voorspelling blijkt dat voor 45 van de 104 opleidingstypen (43%) de voorspelde typering overeenkomt met de realisatie. Dit percentage is lager dan dat voor 2002 (64%), maar hoger dan dat voor 2004 (41%). In totaal waren er in 2007 in 81% van de opleidingstypen een

geheel of bijna juiste typering. Dit is een verbetering is ten opzicht van de 77% in de vorige evaluatie (Dupuy, 2009).

## 5 Conclusies

In dit evaluatierapport zijn specifieke elementen van het prognosemodel geëvalueerd die in 2003 zijn gebruikt voor het opstellen van de uitbreidingsvraag per sector en beroepsgroep en de indicator toekomstige arbeidsmarktsituatie (ITA) voor de periode 2003-2008 (ROA, 2003). De methodiek van de uitbreidingsvraag en de ITA zijn beschreven in *Methodiek arbeidsmarktprognoses en -indicatoren 2003-2008* (Cörvers et al., 2004). Bij de evaluatie in dit rapport werd specifiek ingegaan op welke factoren de afwijkingen tussen de voorspelde en de gerealiseerde uitbreidingsvraag kunnen verklaren. Tevens is de kwaliteit van de prognoses van de uitbreidingsvraag en de ITA voor de periode 2003-2008 zoveel mogelijk vergeleken met de kwaliteit van de eerdere prognoses en ITA. In dit slothoofdstuk wordt in het kort een overzicht gegeven van de belangrijkste bevindingen. De resultaten worden samengevat door deze te vergelijken met die van de kwaliteit van de voorgaande prognoses.

Tabel 5.1 laat het gemiddeld verlies voor de relatieve uitbreidingsvraag zien. Aangezien voor de bedrijfssectorprognoses naast de gegevens van deze evaluatie alleen gegevens van 1998 aanwezig zijn, is het niet mogelijk om iets te zeggen over de ontwikkeling van de uitbreidingsvraagprognoses per sector over de tijd. Als de sectorprognoses van 2007 (prognoses tot 2008) vergeleken worden met die van 1998 dan zien we dat de voorspelkwaliteit verslechterd is. Sterker nog, voor 2007 wordt een score gevonden van 1,19, wat aangeeft dat de SAB-prognoses beter voorspellen dan de sectorprognoses. Met andere woorden, de kwaliteit van de uitbreidingsvraagprognoses per bedrijfssector is achteruit gegaan.

De achteruitgang van de kwaliteit van de uitbreidingsvraagprognoses per sector is niet per definitie terug te brengen op het prognosemodel. Als uitgangspunt voor de bedrijfssectorprognoses zijn de prognoses van het CPB gebruikt. Vervolgens zijn de

procentuele ontwikkelingen en de prognoses van het aantal werkenden weergegeven ten opzicht van het aantal werkenden in het basisjaar 2002 volgens de EBB-data van het CBS. Uit paragraaf 3.3 is echter gebleken dat het veel verschil uitmaakt welke EBB-gegevens gebruikt worden in de analyses (gegevens van 2002 uit 2003 of gegevens van 2002 uit 2009) en of er wel of niet gecorrigeerd wordt voor veranderingen in EBB-data.

**Tabel 5.1**

Het gemiddeld verlies en de score voor de afzonderlijke prognoses

Onderdeel	Gemiddeld verlies (relatief)					Score (relatief)				
	2008	2004	2002	2000	1998	2008	2004	2002	2000	1998
Bedrijfssector										
Uitbreidingsvraag	0,0086				0,0020	1,19				0,52
Beroepsgroep										
Uitbreidingsvraag	0,0295		0,0399	0,0214	0,0229	0,94		0,97	0,77	0,83

Tevens laat Tabel 5.1 zien dat de uitbreidingsvraagprognoses per beroepsgroep verbeterd zijn. Het gemiddeld verlies is 0,0295 in 2007 ten opzichte van 0,0399 in de vorige evaluatie. Ook is de score verbeterd: 0,94 in 2007, ten opzichte van 0,97 in de vorige evaluatie. Met andere woorden, de kwaliteit van de uitbreidingsvraagprognoses per beroepsgroep is vooruit gegaan. De grootste winst in voorspelkwaliteit wordt behaald door rekening te houden met verschuivingen binnen de segmentstructuur, waarvoor het reeds eerder gebruikte random-coëfficiënten-model is gebruikt. Voor de verbijzondering van de sectorgegevens naar beroepssegment is voor de eerste keer het nieuwe model van Cörvers en Dupuy (2003) gebruikt. In het nieuwe model wordt de beroepenstructuur binnen bedrijfssectoren geschat aan de hand van, behalve de tijdtrend, vier factoren, te weten: het productieniveau, de kapitaalintensiteit, het gebruik van nieuwe meer kennisintensieve technologieën en de relatieve lonen op de verschillende arbeidsmarktsegmenten. Door rekening te houden met verschuivingen binnen de beroepenstructuur wordt een lagere score verkregen, hetgeen duidt op een betere voorspelkwaliteit, alhoewel deze verbetering minimaal is.



Tabel 5.2 laat de voorspelkwaliteit van de afzonderlijke prognoses zien. De uitbreidingsvraagprognoses hebben in de meeste gevallen nog steeds een juiste of bijna juiste typering gekregen voor de verschillende beroepsgroepen. 22% van de prognoses heeft een juiste typering tegen over 18% in de vorige evaluatie. Daarnaast zit 57% van de prognoses er maximaal één categorie naast, dit was 61% in de vorige evaluatie. De voorspelkwaliteit van de ITA per opleidingstype is licht verbeterd. Was eerst 41% van de prognoses geheel juist, nu is dat 43%. In de vorige evaluatie zat 77% van de prognoses er maximaal één categorie naast, nu is dat 81%.

**Tabel 5.2**

Voorspelkwaliteit voor de afzonderlijke prognoses

Typering Prognoses tot	Geheel juist (relatief)					Maximaal 1 categorie ernaast (relatief)				
	2008	2004	2002	2000	1998	2008	2004	2002	2000	1998
Beroepsgroep										
Uitbreidingsvraag	22%		18%	42%	40%	57%		61%	68%	66%
Opleidingstype										
ITA	43%	41%	64%	38%	52%	81%	77%	93%	75%	83%

We kunnen concluderen dat de kwaliteit van de uitbreidingsvraagprognoses per sector slechter is dan de prognoses bij de laatste evaluatie. Het verlies is gestegen en de SAB-prognoses voorspellen beter. Wel is gebleken dat het voor de sectorprognoses veel verschil uit kan maken dat er EBB-gegevens gebruikt worden in plaats van Nationale Rekeningen (NR)-data in de analyses, en of er wel of niet gecorrigeerd wordt voor veranderingen in EBB-data.

De kwaliteit van de uitbreidingsvraagprognoses per beroepsgroep is beter dan bij de laatste evaluatie. Het gemiddeld verlies en de score (beide relatief) zijn gedaald ten opzichte van de voorafgaande evaluatie. Het blijkt dat de grootste winst in voorspelkwaliteit wordt behaald door rekening te houden met verschuivingen binnen de segmentstructuur. Dit lijkt er op te wijzen dat het nieuwe model van Cörvers en Dupuy (2003) goed voorspelt, alhoewel de verbetering van de voorspelkwaliteit minimaal is.

De ITA blijkt tenslotte een redelijk goede voorspeller voor de toekomstige arbeidsmarktpositie van een opleiding. Bij opleidingen waarvoor een slecht perspectief werd gegeven, blijken nieuwkomers vaker werkloos te zijn, vaker uit te wijken naar banen buiten de eigen richting, naar deeltijdwerk en tijdelijk werk, en een lager gemiddeld loon te genieten. Met betrekking tot de voorspelkwaliteit van de ITA kan gezegd worden dat deze verbeterd is ten opzichte van de voorafgaande evaluatie. Het blijkt daarnaast verschil uit te maken of de substitutievraag wordt meegenomen in de ITA. De voorspelling verslechtert aanzienlijk als de substitutievraag niet wordt meegenomen.

Tot slot is van belang op te merken dat de beoordeling van de kwaliteit van de prognoses ook afhangt van de wijze waarop geëvalueerd wordt en de data die gebruikt worden bij de evaluatie. Zo blijft het uiteindelijk de vraag hoeveel “beter” de voorspelkwaliteit van de uitbreidingsvraagprognoses per beroepsgroep zou zijn geweest als er betere sectorprognoses waren geweest, immers de sectorprognoses werken ook door in het samenstellen van de prognoses per beroepsgroep. Het is gebleken dat aanpassingen in de EBB achteraf (dus nadat de prognoses gemaakt zijn), van grote invloed kunnen zijn op de evaluatie. Een toename van de afwijking in de EBB-data leidt tot een toename van de voorspelfouten. Verder is gebleken dat het voor de sectorprognoses veel verschil uit kan maken welke EBB-gegevens gebruikt worden in de analyses en of er wel of niet gecorrigeerd wordt voor veranderingen in EBB-data. Bovendien maakt het verschil uit of de prognoses wel of niet gecorrigeerd worden voor veranderingen in EBB-data en of de SAB wel of niet gecorrigeerd wordt door een trend te gebruiken in plaats van de waarde in het basisjaar 2002.

Bij de toekomstige evaluaties van de sectorgegevens is het van belang om de omrekening van CPB-data naar EBB-data goed op elkaar te laten aansluiten. Het nieuwe model van Cörvers en Dupuy (2003), dat gebruikt wordt om de sectorgegevens te verbijzonderen naar beroepssegment, laat een lichte verbetering in voorspelkwaliteit zien. Voor de prognoses tot 2010 is het destijds ontwikkelde

model nog verder ontwikkeld door rekening te houden met interacties tussen beroep-sector-combinaties (Cörvers en Dupuy, 2010). Toekomstige evaluaties zullen moeten uitwijzen of het positieve effect van het nieuwe model blijvend is.

## Literatuur

- Borghans, L. (1993), *Educational Choice and Labour Market Information*, proefschrift, ROA, Universiteit Maastricht.
- Borghans, L. , van Eijs, P. en A. de Grip (1994), *Evaluatie arbeidsmarktprognoses naar opleiding en beroep in 1992*, ROA-R-1994/4, Maastricht.
- Borghans, L., de Grip, A. en J. Hoevenberg (1994), *De bruikbaarheid van deelmarktinformatie voor het informatiesysteem onderwijs-arbeidsmarkt*, ROA-R-1994/13, Maastricht.
- Borghans, L. en J.A.M. Heijke (1994), *Een random-coëfficiënten-model voor het voorspellen van de beroepenstructuur van bedrijfstakken*, ROA-W-1994/1, Maastricht.
- CBS (1993), *Standaard beroepenclassificatie 1992*, Centraal Bureau voor de Statistiek (CBS), Hoofdafdeling Sociale Rekeningen, Sdu/ CBS-publikaties, Den Haag.
- CBS (2010), *Documentatierapport Enquête Beroepsbevolking (EBB) 2009V1, Bijlage 4: Methoden en definities Enquête Beroepsbevolking 2009*, Centraal Bureau voor de Statistiek (CBS), Centrum voor Beleidsstatistiek.
- Cörvers, F. en A. Dupuy (2003), *A micro-economic foundation of modelling and forecasting the occupational structure of economic sectors*, Paper presented at the conference ‘Modeling labour market: Realities and prospects’, Employment Observatory Research Informatics, Athens.
- Cörvers, F., Dupuy, A., Dijksman, S., Goldsteyn, B. en M. Hensen (2004), *Methodiek arbeidsmarktprognoses en- indicatoren 2003-2008*, ROA-W-2004/2, Maastricht.
- Cörvers, F. en A. Dupuy (2006), *Explaining the Occupational Structure of Dutch Sectors of Industry, 1988-2003, 2006/7<sup>E</sup>*, Universiteit Maastricht.

- Cörvers, F. en A. Dupuy (2010), Estimating Employment Dynamics across Occupations and sectors of Industries, *Journal of Macroeconomics*, vol. 32, pp. 17-27.
- CPB (1993), *Athena; Een bedrijfstakkenmodel voor de Nederlandse economie*, Den Haag.
- De Grip, A., Heijke, J.A.M. en H. Berendsen (1991), *Eerste evaluatie informatiesysteem onderwijs-arbeidsmarkt*, ROA-R-1991/1, Maastricht.
- Dupuy, A. (2005), *An evaluation of labour market forecasts by type of education and occupation for 2002*, ROA-W-2005/1E, Maastricht.
- Dupuy, A. (2009), *An evaluation of the forecast of the indicator of the labour market gap*, ROA-TR-2009/3, Maastricht.
- Engle, R.F. en C.W.J. Granger (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, vol. 55, issue 2, pp. 251-276.
- Granger, C.W.J. en P. Newbold (1986), *Forecasting Economic Time Series*, Orlando.
- ROA (2002), *ROA-classificatiegids 2002*, ROA-R-2002/3, Maastricht.
- ROA (2003), *De arbeidsmarkt naar opleiding en beroep tot 2008*, ROA-R-2003/11, Maastricht.
- ROA (2009), *De arbeidsmarkt naar opleiding en beroep tot 2014*, ROA-R-2009/5, Maastricht.
- Smits, W., en B. Diephuis (2001), *Evaluatie arbeidsmarktprognoses naar opleiding en beroep tot 1998*, ROA-R-2001/2, Maastricht.
- Theil, H. (1958). *Economic Forecasts and Policy*, Amsterdam.

## Bijlage 1

**Tabel B.1**

Vergelijking verlies uitbreidingsvraagprognoses per bedrijfssector, ROA rapport en CPB<sup>19</sup>

	Referentie	Prognose	Realisatie	Voorspelfout	Verlies	Gemiddeld
<i>Bedrijfssectoren (ROA) (zie Tabel 3.1)</i>	2002	2007	2007			Verlies
Landbouw en visserij	226	174	214	40	0,0310	0,0010
Voeding	139	142	133	-9	0,0044	0,0001
Chemie	108	104	129	25	0,0521	0,0008
Metaal en elektrotechniek	380	356	396	40	0,0111	0,0006
Overige industrie	357	339	292	-47	0,0175	0,0009
Energie	53	58	58	0	0,0002	0,0000
Bouw en onroerend goed	544	524	570	46	0,0072	0,0006
Handel en reparatie	1.051	1.062	946	-117	0,0123	0,0018
Transport en communicatie	412	395	448	53	0,0167	0,0010
Bank- en verzekeringswezen	254	287	252	-35	0,0190	0,0007
Horeca en zakelijke dienstv.	1.178	1.228	1.217	-11	0,0001	0,0000
Kwartaire diensten	1.161	1.266	1.299	33	0,0008	0,0001
Overheid en onderwijs	976	1.089	1.010	-79	0,0065	0,0009
Bedrijfssectoren totaal	7.035	7.224	7.259	35	0,0000	0,0085

	Referentie	Prognose	Realisatie	Voorspelfout	Verlies	Gemiddeld
<i>Bedrijfssectoren CPB</i>	2002	2007	2007			Verlies
Landbouw en visserij	249	232	240	8	0,0011	0,0000
Voedings- en genotsmiddelenindustrie	142	140	140	1	0,0000	0,0000
Chemische, rubber- en kunststofverw. ind.	109	102	108	6	0,0030	0,0000
Metalektro-industrie	389	381	370	-11	0,0009	0,0000
Overige industrie	364	361	343	-17	0,0023	0,0001
Delfstofwinning	8	9	10	1	0,0147	0,0000
Openbare nutsbedrijven	38	36	33	-4	0,0103	0,0001
Aardolie-industrie	7	8	6	-1	0,0372	0,0000
Bouwnijverheid en – installatiebedrijven	493	522	484	-37	0,0057	0,0004
Verhuur van en handel in onroerend goed	58	59	73	14	0,0610	0,0005
Handel en reparatiebedrijven	1.186	1.240	1.162	-78	0,0044	0,0007
Transport- en opslagbedrijven	330	341	332	-9	0,0007	0,0000
Communicatiebedrijven	114	117	115	-3	0,0006	0,0000
Bank- en verzekeringswezen	299	319	279	-39	0,0173	0,0007
Uitzendbureaus en huishoudelijke diensten	467	467	404	-62	0,0179	0,0012
Overige tertiaire diensten	1.255	1.347	1.365	18	0,0002	0,0000
Gezondheids- en welzijnszorg	870	939	916	-23	0,0007	0,0001
Overheid	852	887	873	-14	0,0003	0,0000
Bedrijfssectoren totaal	7.231	7.505	7.258	-247	0,0012	0,0040

19. Referentie 2002, prognose 2007, realisatie 2007 en voorspelfout in duizendtallen.

**Tabel B.2**

Totaaloverzicht voorspelkwaliteit uitbreidingsvraagprognoses per bedrijfssector

			Gem. verlies prognose	Gem. verlies SAB	Score
<i>EBB-data</i>	Prognoses	2007			
<i>van het jaar 2002</i>	<i>Absoluut</i>		0,0085	0,0074	1,15
<i>gegevens uit 2009</i>	<i>Relatief</i>		0,0086	0,0073	1,19
(Zie Tabel 3.2)					
<i>EBB-data</i>	Prognoses	2007			
<i>van het jaar 2002</i>	<i>Absoluut</i>		0,0061	0,0060	1,01
<i>gegevens uit 2003</i>	<i>Relatief</i>		0,0085	0,0061	1,38

**Tabel B.3**

Relatie tussen voorspelde arbeidsmarktpositie en huidige arbeidsmarktsituatie

		Indicator arbeidsmarktsituatie	$\alpha$	T-waarde	$\beta$	T-waarde	
<i>ITA</i> (zie ook Tabel 4.1)	$\Delta Gap_j$	-7,003	-3,45	***	27,384	3,85	**
	$Gap_{j,08}$	49,187	17,21	***	48,656	3,05	***
<i>ITA zonder kortdurig werklozen</i>	$\Delta Gap_j$	-5,476	-2,53	**	27,210	1,97	*
	$Gap_{j,08}$	51,767	17,28	***	43,758	2,47	**
<i>ITA zonder substitutievraag</i>	$\Delta Gap_j$	-7,284	-3,59	***	32,131	2,36	**
	$Gap_{j,08}$	48,817	16,90	***	48,764	2,84	**

\*\*\* significant bij betrouwbaarheid van 99%

\*\* significant bij betrouwbaarheid van 95%

\* significant bij betrouwbaarheid van 90%

**Tabel B.4**

Relatie tussen voorspelde arbeidsmarktpositie en huidige arbeidsmarktsituatie

	Indicator arbeidsmarktsituatie	$\alpha$	T- waarde		$\beta$	T-waarde	
<i>ITA</i>							
(zie ook Tabel 4.2)	<i>Werkloos</i>	0,948	44,90	***	0,021	4,84	***
	<i>&gt; 4 maanden werkloos</i>	0,999	34,91	***	0,002	0,70	
	<i>Onderbenutting</i>	0,985	33,23	***	0,001	1,31	
	<i>Buiten eigen richting</i>	0,912	32,70	***	0,004	4,42	***
	<i>Deeltijdwerk</i>	0,953	32,34	***	0,002	2,78	***
	<i>Vast werk</i>	1,253	19,46	***	-0,004	-3,79	***
	<i>Gemiddeld loon</i>	1,285	27,17	***	-0,000	-5,99	***
<i>ITA zonder kortdurig werklozen</i>							
	<i>Werkloos</i>	0,906	45,77	***	0,016	3,90	***
	<i>&gt; 4 maanden werkloos</i>	0,959	36,77	***	0,000	-0,08	
	<i>Onderbenutting</i>	0,939	34,73	***	0,001	0,89	
	<i>Buiten eigen richting</i>	0,877	33,69	***	0,003	3,71	***
	<i>Deeltijdwerk</i>	0,917	33,67	***	0,001	1,98	*
	<i>Vast werk</i>	1,122	18,69	***	-0,003	-2,83	***
	<i>Gemiddeld loon</i>	1,186	27,04	***	-0,000	-5,48	***
<i>ITA zonder substitutievraag</i>							
	<i>Werkloos</i>	0,957	49,18	***	0,02	5,11	***
	<i>&gt; 4 maanden werkloos</i>	0,986	37,44	***	0,005	1,72	*
	<i>Onderbenutting</i>	0,993	35,99	***	0,001	1,38	
	<i>Buiten eigen richting</i>	0,936	35,38	***	0,003	3,91	***
	<i>Deeltijdwerk</i>	0,973	35,10	***	0,002	2,37	**
	<i>Vast werk</i>	1,253	21,04	***	-0,004	-3,99	***
	<i>Gemiddeld loon</i>	1,222	26,17	***	-0,000	-4,49	***

\*\*\* significant bij betrouwbaarheid van 99%

\*\* significant bij betrouwbaarheid van 95%

\* significant bij betrouwbaarheid van 90%

**Tabel B.5**

Kwalitatieve typering van de ITA per onderwijstype

		<i>Realisatie</i>					
<i>ITA (zie ook Tabel 4.4)</i>	<i>Prognose</i>	Zeer goed	Goed	Redelijk	Matig	Slecht	<i>Totaal</i>
	Zeer goed	0	5	3	0	0	8
	Goed	1	24	10	5	6	46
	Redelijk	0	2	1	5	1	9
	Matig	0	2	9	11	4	26
	Slecht	0	1	2	3	9	15
	<i>Totaal</i>	1	34	25	24	20	104
<i>ITA zonder kortdurig werklozen</i>	Zeer goed	0	11	0	0	2	13
	Goed	1	38	10	1	5	55
	Redelijk	0	2	3	3	2	10
	Matig	0	11	2	2	2	17
	Slecht	0	2	0	7	0	9
	<i>Totaal</i>	1	64	15	13	11	104
<i>ITA zonder substitutievraag</i>	Zeer goed	0	2	1	1	0	4
	Goed	1	15	15	7	5	43
	Redelijk	0	6	1	6	1	14
	Matig	0	4	8	9	4	25
	Slecht	0	1	4	7	6	18
	<i>Totaal</i>	1	28	29	30	16	104